



HAL
open science

Iterative predistortion algorithms adapted to the increasing throughput of satellite communications

Nicolas Alibert

► **To cite this version:**

Nicolas Alibert. Iterative predistortion algorithms adapted to the increasing throughput of satellite communications. Signal and Image processing. Ecole nationale supérieure Mines-Télécom Atlantique, 2019. English. NNT: 2019IMTA0126 . tel-02096912

HAL Id: tel-02096912

<https://theses.hal.science/tel-02096912>

Submitted on 11 Apr 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

THESE DE DOCTORAT DE

L'ÉCOLE NATIONALE SUPERIEURE MINES-TELECOM ATLANTIQUE
BRETAGNE PAYS DE LA LOIRE - IMT ATLANTIQUE
COMUE UNIVERSITE BRETAGNE LOIRE

ECOLE DOCTORALE N° 601
*Mathématiques et Sciences et Technologies
de l'Information et de la Communication*
Spécialité : *Télécommunications*

Par

Nicolas ALIBERT

Algorithmes itératifs de prédistorsion adaptés à la montée en débit des communications par satellite

Thèse présentée et soutenue à l'IMT Atlantique – Campus de Brest, le 25 Janvier 2019

Unité de recherche : IMT Atlantique, Lab-STICC, UBL, F-29238 Brest, France

Thèse N° : 2019IMTA0126

Rapporteurs avant soutenance :

Geneviève BAUDOIN, Professeur, ESIEE Paris
Charly POUILLIAT, Professeur, INP-ENSEEIH T Toulouse

Composition du Jury :

Président : Daniel ROVIRAS, Professeur, CNAM

Examineurs : Mathieu DERVIN, Ingénieur de recherche, Thales Alenia Space
Damien CASTELAIN, Chief Scientist, Mitsubishi Electric R&D Centre Europe
Charlotte LANGLAIS, Maître de conférences, IMT Atlantique Bretagne - Pays de la Loire

Rapporteurs : Geneviève BAUDOIN, Professeur, ESIEE Paris
Charly POUILLIAT, Professeur, INP-ENSEEIH T Toulouse

Dir. de thèse : Karine AMIS, Professeur, IMT Atlantique Bretagne - Pays de la Loire

Remerciements

“The art of life is more like the wrestler’s art than the dancer’s, in respect of this, that it should stand ready and firm to meet onsets that are sudden and unexpected.”

Marcus Aurelius - Meditations, VII.

Je tiens à remercier Karine Amis, ma directrice de thèse, pour son aide tout au long de la thèse. Sans son investissement, sa patience et ses conseils pendant ces trois années, je n’aurais pas appris autant et cette thèse n’aurait pas été la même. Son ouverture d’esprit, et la confiance qu’elle m’a accordé lorsque je lui ai présenté des idées ont rendu très agréable de travailler avec elle.

De même, je veux remercier mes encadrants, Charlotte Langlais et Damien Castelain, qui ont été là pour me conseiller, pour les échanges scientifiques très enrichissantes que j’ai pu avoir avec eux et pour la grande autonomie qu’ils m’ont laissé. Ces derniers ont permis d’avoir une autre perspective sur les problèmes que j’ai pu rencontrer.

Je tiens aussi à remercier le président du jury, Daniel Roviras, les rapporteurs, Geneviève Baudoin et Charly Pouillat, et les membres du jury, Mathieu Dervin, Charlotte Langlais, Karine Amis et Damien Castelain pour le temps consacré et les échanges enrichissants lors de la soutenance.

Ainsi que les membres de mon comité de suivi individuel, Ammar Sharaiha et Yves Louët, pour leurs conseils et leur bienveillance. Je tiens à remercier plus particulièrement Yves Louët qui m’a donné le goût pour les communications numériques lors de mon cursus d’ingénieur.

Je veux remercier les collègues de Mitsubishi Electric R&D Centre Europe pour leur accueil et leurs gentilleses.

La thèse n’aurait pas été la même sans mes chers camarades doctorants du département SC et des autres départements.

Et enfin, je tiens à remercier ma famille pour leur soutien.

Contents

Remerciements	3
Acronyms	9
1 Introduction	15
1.1 Motivation	15
1.1.1 Satellite communication links	15
1.1.2 Broadband and broadcast Satellite	16
1.2 Problem	17
1.3 Thesis Outline	20
1.4 List of publications	20
2 Communications over satellite channels	23
2.1 Introduction	24
2.2 Satellite transmission	24
2.3 Specificities of the DVB-S2(X) and beyond	26
2.3.1 Constellation	26
2.3.2 Pulse shaping filter	27
2.3.3 SC-OFDM	27
2.4 Description of the transponder model	28
2.4.1 IMUX and OMUX filters	28
2.4.2 Amplifiers	29
2.5 Modelling of the satellite transponder	33
2.5.1 Wiener-Hammerstein	33
2.5.2 Volterra model	34
2.5.3 Memory Polynomial	36
2.5.4 Identification	37
2.6 Impairments due to the satellite transponder	37
2.6.1 Wrapping	39
2.6.2 Clustering	39
2.6.3 Spectral regrowth	41
2.7 Performance metrics	42
2.7.1 Total degradation	42
2.7.2 NMSE	43
2.7.3 ACI factor	45
2.8 Conclusion	45

3	State-of-the-art of predistortion techniques	47
3.1	Introduction	48
3.2	Direct predistortion method	48
3.2.1	P-th order inverse	48
3.2.2	Global compensation	49
3.2.3	Learning architecture	50
3.2.4	Neural network as a predistortion technique	53
3.3	Iterative predistortion methods	54
3.3.1	Small variation algorithm	55
3.3.2	Iterative predistortion based on the fixed-point theorem	56
3.3.3	Successive approximation	57
3.4	Conclusion	59
4	Iterative predistortion	61
4.1	Introduction	62
4.2	Data-optimized signal predistortion	62
4.2.1	Data predistortion based on the contraction mapping theorem	62
4.2.2	Data-optimized signal predistortion scheme	64
4.2.3	Comparison of the data-optimized signal predistortion with the successive approximation scheme assuming perfect identification	64
4.2.4	Comparison with signal predistortion based on the contraction mapping	66
4.3	Iterative adaptive signal predistortion	69
4.3.1	Iterative adaptive signal-optimized predistortion	70
4.3.2	Iterative adaptive data-optimized signal predistortion	70
4.3.3	Adaptation process	71
4.3.4	Comparison with signal-optimized signal predistortion scheme	74
4.3.5	Robustness towards feedback mismatch	76
4.4	Perspectives	76
4.4.1	Backpropagation	76
4.4.2	Joint Clipping and predistortion	79
4.4.3	Variable step-size	79
4.5	Conclusion	80
5	Conclusion and perspectives	83
5.1	Conclusion	83
5.2	Perspectives	84
5.2.1	Similarity with Neural Network	84
5.2.2	Use of piecewise-defined function	85
5.2.3	Iterative processing for FTN	85
5.2.4	Joint adaptation of the constellation and predistortion	85
	Appendices	87

A Derivation of the conditional variance and pseudo-variance for a third order Volterra model	87
A.1 Mathematical foreword	88
A.2 Conditional variance for a third order Volterra model	88
A.3 Conditional pseudo-variance for a third order Volterra model	89
B Computation of the derivative of Saleh's HPA model	91
B.1 Derivative of Saleh's HPA model	91
C Résumé	93
C.1 Introduction et contexte	93
C.2 Communications sur un lien satellitaire	96
C.3 Etat de l'art des méthodes de prédistorsion	97
C.4 Prédistorsion itérative	98
C.5 Conclusion et perspectives	98

Acronyms

3GPP Third Generation Partnership Project.

5G fifth generation of mobile telecommunications standard.

ACI Adjacent Carrier Interference.

AM-AM Amplitude-to-amplitude Modulation.

AM-PM Amplitude-to-phase Modulation.

APSK Amplitude and Phase Shift Keying.

AWGN Additive White Gaussian Noise.

BIBO Bounded Input Bounded Output.

BPSK Binary Phase Shift Keying.

DC Direct Current.

DFT Discrete Fourier Transform.

DLA Direct Learning Architecture.

DTH Direct-To-Home.

DVB-S2 Digital Video Broadcasting - Satellite - Second Generation.

DVB-S2X Digital Video Broadcasting - Satellite - Second Generation Extension.

ESA European Space Agency.

F ϵ Filtered- ϵ .

FTN Faster-than-Nyquist.

F x Filtered-X.

HPA High Power Amplifier.

IBO Input Back-Off.

- IEL** Instantaneous Equivalent Linear Filter.
- ILA** Indirect Learning Architecture.
- IMUX** Input MUltipleXer.
- INP** Instantaneous Normalized Power.
- LEO** Low earth orbit.
- LLR** Log-Likelihood Ration.
- LMS** Least Mean Squares.
- LTE** Long Term Evolution.
- LUT** Look-Up Table.
- MEO** Medium earth orbit.
- MLP** MultiLayer Perceptron.
- MODCOD** MODulation and CODing.
- NFX** Non-linear filtered-X.
- NFXLMS** Non-linear filtered-X [Least Mean Squares \(LMS\)](#).
- NISI** Non-linear Inter-Symbol Interference.
- NMSE** Normalised Mean Square Error.
- OBO** Output Back-Off.
- OFDM** Orthogonal Frequency Division Multiplexing.
- OFDMA** Orthogonal Frequency Division Multiple Access.
- OMUX** Output MUltipleXer.
- PAPR** Peak-to-Average Ratio.
- PSD** Power Spectral Density.
- PSK** Phase-shift keying.
- QAM** Quadrature Amplitude Modulation.
- QPSK** Quadrature Phase Shift Keying.
- RF** Radio Frequency.
- RLS** Recursive Least Squares.
- SC-FDMA** Single-Carrier Frequency Division Multiple Access.

SC-OFDM Single Carrier Orthogonal Frequency Division Multiplexing.

SINR Signal-to-Interference plus Noise Ratio.

SIR Signal-to-Interference Ratio.

SNR Signal-to-Noise Ratio.

SRRC Square Root Raised Cosine.

SSPA Solid State Power Amplifier.

TD Total Degradation.

TWTA Travelling-Wave Tube Amplifier.

UMTS Universal Mobile Telecommunications System.

VL-SNR Very-Low Signal-to-Noise Ratio.

List of Figures

1.1	Direct-to-home topology	17
1.2	Constellation received for different IBOs - 32-APSK - SRRC $\beta = 5\%$	18
1.3	Constellation received for different modulation - $IBO = 12\text{dB}$ - SRRC $\beta = 5\%$	18
1.4	Constellation received for different roll-offs - 32-APSK - $IBO = 12\text{dB}$	19
2.1	Communication Chain	25
2.2	Illustration of the transmission model at signal and symbol rate .	25
2.3	32-APSK [11].	26
2.4	Complementary cumulative distribution function of the Instantaneous Normalized Power for different pulse shaping filters - QPSK	27
2.5	SC-OFDM transmitter and receiver	28
2.6	IMUX and OMUX characteristics [11].	29
2.7	AM-AM and AM-PM characteristics proposed in the DVB-S2 standard [11].	30
2.8	Amplitude-to-amplitude Modulation (AM-AM) and Amplitude-to-phase Modulation (AM-PM) of the Saleh model (2,1,2,1.5) and its polynomial approximation	32
2.9	Illustration of IBO and OBO over the Saleh model (2,1,2,1.5) . .	33
2.10	Wiener-Hammerstein model	34
2.11	Identification principle	37
2.12	Clustering and wrapping for 16-APSK – Pulse shaping: $\beta = 0.05$ – $IBO = 12\text{dB}$ - Noiseless without IMUX and OMUX filters . . .	38
2.13	Conditional variance for different phase offset of the QPSK. . . .	42
2.14	2+4+2 APSK for different phase offset of the QPSK and different Saleh model.	43
2.15	Power spectral density over frequency for different Input Back-Off (IBO)	44
2.16	Power spectral density over frequency for different orders of approximation of the function	44
3.1	Global compensation - Dynamic version	50
3.2	Indirect Learning Architecture	51
3.3	Direct Learning Architecture	53
3.4	Small variation update mechanism	55
3.5	k -th stage of the iterative signal predistortion scheme.	58

3.6	Successive approximation - First scheme - Suppression of the spectral regrowth [30]	59
3.7	Successive approximation - Second scheme - Suppression of the in-band distortion [30]	60
3.8	Successive approximation - Third scheme - Combination of the two scheme (suppression of spectral regrowth and in-band distortion)[30]	60
4.1	Data predistortion structure	63
4.2	Total degradation for a target BER of 10^{-4}	65
4.3	NMSE at the output of the matched filter for an IBO of 14 dB versus the number of stages	66
4.4	ACI at the output of the matched filter versus the OBO.	67
4.5	NMSE at the output of the matched filter for an IBO of 14 dB versus the number of stages.	67
4.6	Total degradation with ACI for a target BER of 10^{-4} for the data-optimized and the signal-optimized signal predistortion scheme with perfect knowledge of the channel (Ideal) or with identification (MP).	68
4.7	First stage of the proposed iterative adaptive signal predistortion scheme.	69
4.8	k -th stage ($k \geq 1$) of the proposed iterative adaptive signal predistortion scheme.	69
4.9	First stage of the data-optimized iterative adaptive signal predistortion scheme.	70
4.10	k -th stage ($k \geq 1$) of the data-optimized iterative adaptive signal predistortion scheme.	71
4.11	Adaptation of the k -th stage.	72
4.12	ANMSE versus the number of frames with and without pregain.	74
4.13	ACI at the output of the matched filter versus the OBO.	74
4.14	NMSE versus the OBO.	75
4.15	Total degradation with ACI for a target BER of 10^{-4} for the reference method and the proposed method for 10 iterations.	75
4.16	AMSE versus number of frames p	75
4.17	Total degradation with ACI for a target BER of 10^{-4} for different feedback approximation.	77

Chapter 1

Introduction

“Don’t be so gloomy. After all it’s not that awful. You know what the fellow said – in Italy, for thirty years under the Borgias, they had warfare, terror, murder and bloodshed, but they produced Michelangelo, Leonardo da Vinci and the Renaissance. In Switzerland, they had brotherly love, they had five hundred years of democracy and peace – and what did that produce? The cuckoo clock.”

Orson Welles as Harry Lime.
The Third Man. Directed by Carol Reed.
British Lion Film Corporation. 1949

Contents

1.1 Motivation	15
1.1.1 Satellite communication links	15
1.1.2 Broadband and broadcast Satellite	16
1.2 Problem	17
1.3 Thesis Outline	20
1.4 List of publications	20

1.1 Motivation

1.1.1 Satellite communication links

Satellite communications appeared after the end of the Second World War. The development of long-range rockets during the war, and notably the V2, by the Germans, enabled to start the space conquest and, therefore, to launch artificial satellites [1].

The principal benefit of using extra terrestrial relays over any ground-based communication system is ability to cover wide areas. To the point that global

coverage can be provided by only three geostationary satellites [1]. Then, satellite communication links are competitive for broadcast purpose.

Another important advantage of satellite-based communication is its robustness towards technological and natural disaster. For instance, the program from the [European Space Agency \(ESA\)](#), called EMERGSAT, completed in 2008, aimed to provide satellite communication links between multiple fixed or portable and even moving stations within a few hours over Europe and the Mediterranean basin after a disaster [2].

The unneeded deployment of heavy ground-based infrastructure for backhauling of mobile networks in order to cover un-served or under-served areas is also an advantage to deploy satellite networks in developing countries [3].

With the reduction of the satellite launch cost, due to partial reuse and bigger rocket launcher, huge constellations of small satellites are investigated and/or planned to be launched by multiple private companies (in [Low earth orbit \(LEO\)](#), [Medium earth orbit \(MEO\)](#)) in order to provide a low-latency global internet connectivity [4], [5].

For airborne, high-speed trains and maritime Wi-Fi services, satellite-provided service seems to be a good alternative to ground-based one. It is easier to cover large distance and bodies of water, remote terrain and rural area through satellite communication links [6]–[9].

Those advantages, the ability to provide coverage everywhere and more easily to maritime, airborne and railway and the potential for backhauling which may ease the scalability of [fifth generation of mobile telecommunications standard \(5G\)](#) network, are pushing the integration of a satellite part into the 5G [9].

All the previously mentioned applications lead to a need for higher data rates over satellite communications links which involves enhanced technological solutions.

1.1.2 Broadband and broadcast Satellite

In the case of video broadcasting, the introduction of high-definition and ultra-high-definition television (UHDTV) and satellite news gathering (SNG) mainly account for the need of more throughput.

The DVB-S was first released in 1994, and was mainly designed for direct-to-home television broadcast via satellite. It doesn't offer much flexibility in terms of spectral efficiency since it is constrained by the use of one modulation, the QPSK, and an unique roll-off of 35% [10].

This issue leads to the development of the DVB-S2, released in 2005. To provide higher capacity, higher modulation order (from 8-PSK up to 32-APSK) and sharper roll-off (up to 20%) were introduced [11].

The extension of the DVB-S2 standard, the DVB-S2X, was approved, in 2014. It enables to operate at [Very-Low Signal-to-Noise Ratio \(VL-SNR\)](#) and very high signal-to-noise ratio and increases the granularity of the [MODulation and CODing \(MODCOD\)](#) available. In order to do so, the [Digital Video Broadcasting - Satellite - Second Generation Extension \(DVB-S2X\)](#) added, among other solutions, even higher order modulation (from 64-APSK up to 256-APSK) and even more sharper roll-off (up to 5%) [12].

For broadband or broadcast satellite, which are generally *direct-to-home* applications, the star topology is often used [13]. It is a point-to-multipoint connectivity. In that case, the transmitter (also called hub, gateway or feeder station)

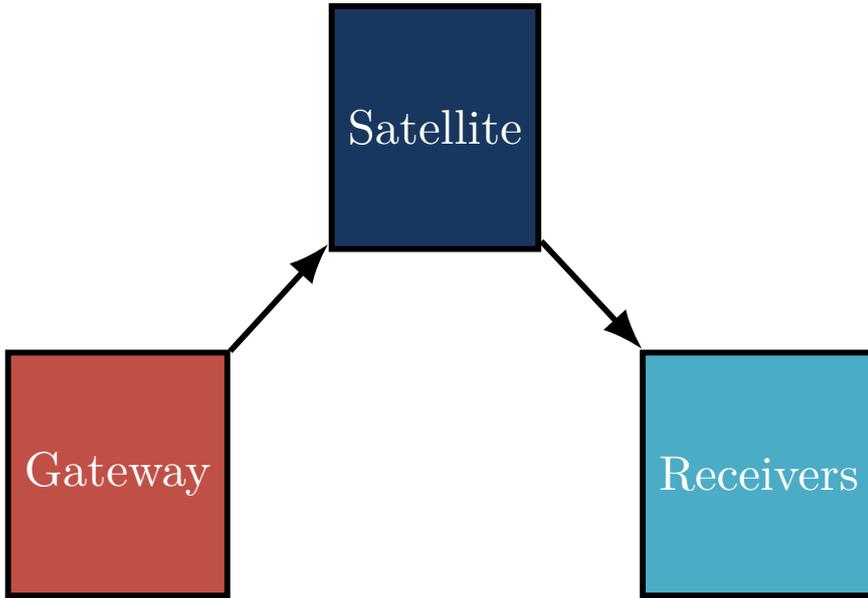


Figure 1.1: Direct-to-home topology

sends a signal to the satellite, then the satellite transponder amplifies and often frequency shifts the received signal and retransmits towards the ground-based receivers. Figure 1.1 illustrates this scheme.

Broadcast satellite system consists of a gateway which sends data to end-users via satellite. The link from the gateway to the satellite is referred to as *uplink* while the link from the satellite to the end-users is referred to as *downlink*.

The main source of impairment over this type of link is due to the satellite transponder. The on-board amplifier response tends to be nonlinear, which causes **Non-linear Inter-Symbol Interference (NISI)**. The distortion can only be avoided at the cost of an increased **Output Back-Off (OBO)**, which reduces the power available for the receivers but allows the amplifier to have a more linear characteristic. Another solution could be to oversize the dimension of the amplifier to provide sufficient power. However, the power and weight constraints imposed to the satellite load prevent this solution.

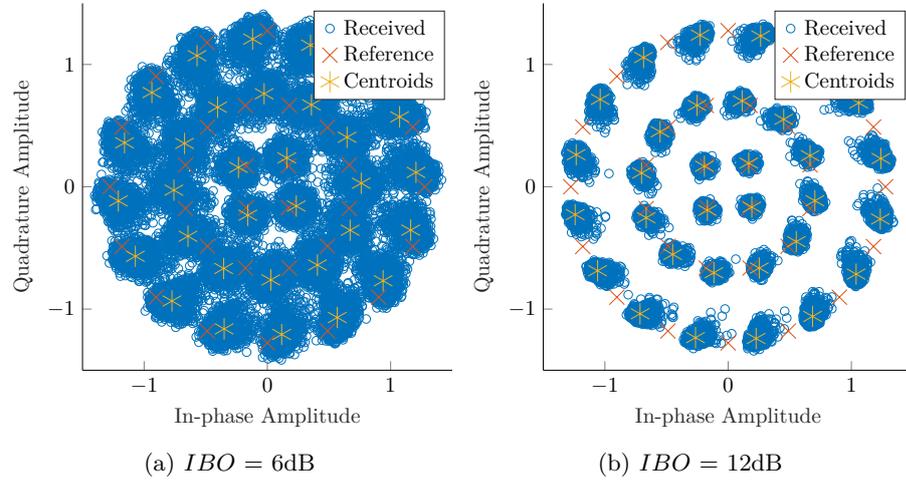
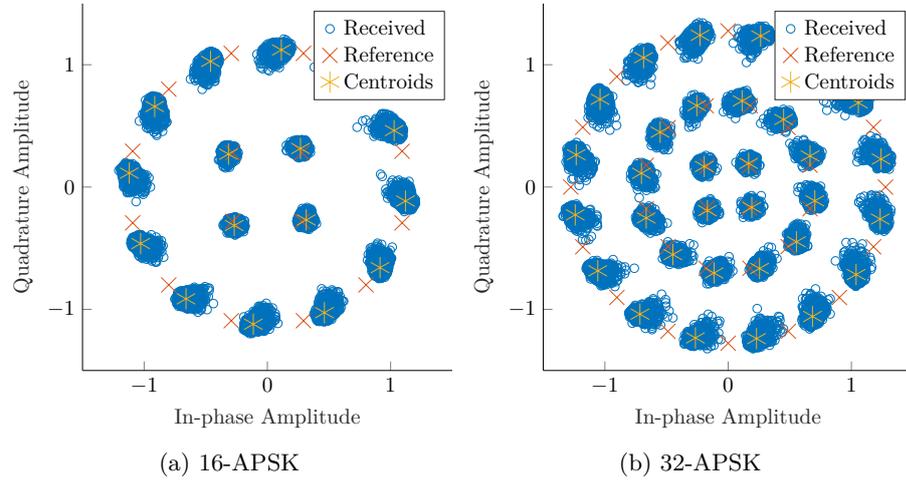
Nonlinear distortion is further increased with the use of higher order modulation and sharper roll-off since they increase the **Peak-to-Average Ratio (PAPR)** and therefore drive more the amplifier into distortion.

Figures 1.2 to 1.4, respectively show that the smaller the **IBO**, the higher the modulation order and sharper the roll-off, the worst the degradation due to distortion.

1.2 Problem

As we have seen, the classic technique to increase the throughput such as using higher modulation order and sharper roll-off tends to increase the **NISI**.

Therefore, in order to increase the throughput the mitigation of the nonlinearity is a key enabling technology.

Figure 1.2: Constellation received for different IBOs - 32-APSK - SRRC $\beta = 5\%$ Figure 1.3: Constellation received for different modulation - $IBO = 12\text{dB}$ - SRRC $\beta = 5\%$

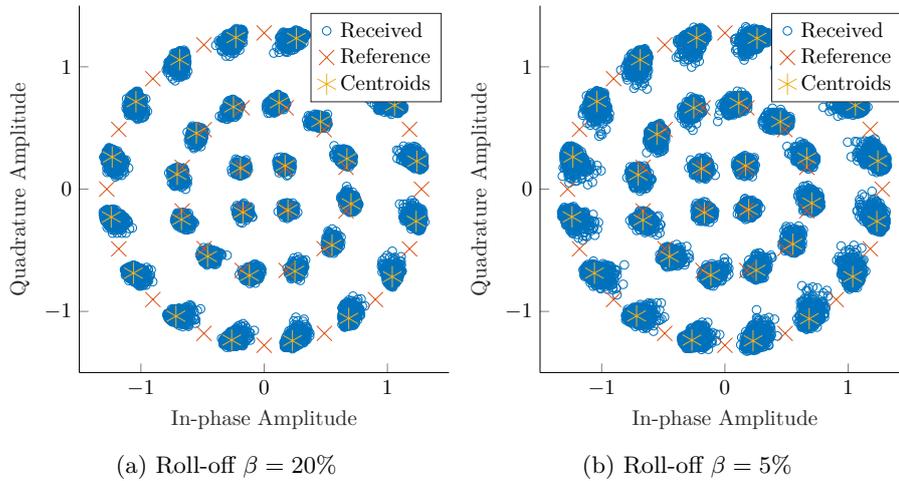


Figure 1.4: Constellation received for different roll-offs - 32-APSK - $IBO = 12\text{dB}$

There are multiple ways to tackle this problem, among the most commonly found in the literature, one can first cite reduction of the crest factor, predistortion, on-board linearisation and equalization, for instance as in [14]–[17].

The reduction of the crest factor is often applied with multi-carrier systems rather than single-carrier systems as the crest factor is less important with the latter.

In a broadcast context, the receivers are cost-driven. Then, their processing capability should be kept to a minimum. Moreover, nonlinear equalization might enhance noise.

The on-board linearisation might be one of the most effective method since it only has to deal with the nonlinear amplifier response and not with the whole transponder response. As a consequence, the memory effects induced by the filtering before and after are not to be taken into account. However, those solutions request to increase the payload and therefore also increase the launching cost of the satellite.

As a consequence, the predistortion is often a good way to mitigate nonlinearities while keeping the cost of receivers and satellites low. The computation power available at the gateway is often large enough to perform predistortion operation.

In the literature, most of the predistortion techniques, the performance of the proposed method relies on the chosen structure for the predistortion. Other methods rely only on the chosen structure for the identification of the nonlinear characteristics of the satellite for which we have more prior information.

Iterative methods seems to outperform direct method, however, those solutions often relies on empirical tuning of parameters.

This PhD focuses on nonlinear distortion management for [Digital Video Broadcasting - Satellite - Second Generation \(DVB-S2\)](#) communication. In a first attempt, to circumvent the issue of both requiring to define a predistorter structure and adapt empirical tuning, we propose a iterative method based on the contraction mapping aiming at minimizing the [Normalised Mean Square](#)

Error (NMSE). The proposed iterative method relies only on the identification of the channel. Therefore, the performances are only depend on the quality of the identification and not the choice of the structure of the predistorter.

However, the verification of hypothesis of contraction mapping are norm-dependent. Moreover, the best performance in terms of **Total Degradation (TD)** are obtained when the trade-off between the mitigation of the distortion and the **OBO** is achieved. Therefore in this case, some distortions are still remaining which is, therefore, not in the scope of the theorem. Nevertheless, performing a few iteration might still improve the linearisation performance [18].

Another issue arise when identification is used, the mismatch between the structure requires to introduce an empirical tuning.

As a consequence, we have introduced tuning parameters and proposed an adaptation method which allows to ensure the improvement of the linearisation performance.

1.3 Thesis Outline

The manuscript is organised as follows:

Chapter 2 In this chapter, we define the satellite communications channel and the physical layer of the **DVB-S2X**. Then, we discuss the different impairments generated by the satellite transponder.

Chapter 3 We present the predistortion solution proposed in the literature to counter the effect of distortion introduces the transponder.

Chapter 4 Building on Chapter 3, we propose iterative predistortion scheme in order to mitigate the nonlinearities.

Chapter 5 This chapter contains some final thoughts and perspectives to investigate in future research.

1.4 List of publications

International conference papers

- N. Alibert, K. Amis, C. Langlais, D. Castelain. "Iterative adaptive signal predistortion for satellite communications", ISTC 2018: International Symposium on Turbo Codes & Iterative Information Processing 2018, Hong Kong, Hong Kong.
- N. Alibert, K. Amis, C. Langlais, D. Castelain. "Comparison of signal predistortion schemes based on the contraction mapping for satellite communications with channel identification", ICT 2018: 25th International Conference on Telecommunications, Saint Malo, France.

National conference papers

- N. Alibert, K. Amis, C. Langlais, D. Castelain. "Signal predistortion scheme based on the contraction mapping theorem", GRETSI 2017 : 26ème colloque du Groupement de Recherche en Traitement du Signal et des Images, 2017, Juan-les-Pins, France.

Presentation

- N. Alibert, D. Castelain. "Correction of the non-linearity at the transmitter", TM-S0403, DVB TM-S GFT meeting, 7th of November 2017

Chapter 2

Communications over satellite channels

“A transmission? Out here?”

Sigourney Weaver as Ellen Ripley.
Alien. Directed by Ridley Scott.
20th Century Fox. 1979

Contents

2.1	Introduction	24
2.2	Satellite transmission	24
2.3	Specificities of the DVB-S2(X) and beyond	26
2.3.1	Constellation	26
2.3.2	Pulse shaping filter	27
2.3.3	SC-OFDM	27
2.4	Description of the transponder model	28
2.4.1	IMUX and OMUX filters	28
2.4.2	Amplifiers	29
2.5	Modelling of the satellite transponder	33
2.5.1	Wiener-Hammerstein	33
2.5.2	Volterra model	34
2.5.3	Memory Polynomial	36
2.5.4	Identification	37
2.6	Impairments due to the satellite transponder	37
2.6.1	Wrapping	39
2.6.2	Clustering	39
2.6.3	Spectral regrowth	41
2.7	Performance metrics	42
2.7.1	Total degradation	42
2.7.2	NMSE	43
2.7.3	ACI factor	45
2.8	Conclusion	45

2.1 Introduction

In this chapter, we first present a general satellite transmission scheme which is similar to the reference DVB-SX [Direct-To-Home \(DTH\)](#) simulation scenario. Secondly, we detail the specificities of the DVB-S2(X), which standardise the physical layer of the transmitter. Therefore, we characterize the constellation and waveform standardized in the DVB-S2(X) as well as a candidate for future standard. Then, we describe the satellite transponder, its filters and different types of amplifiers that can be found in a satellite transponder. Subsequently, we present mathematical models for the satellite transponder that can be used for predistortion design and channel identification. Given one transponder model, we illustrate the different impairments generated by the satellite transponder. Then, we discuss the performance metrics commonly used to measure the performance of the countermeasures to impairments.

2.2 Satellite transmission

Let us now describe the baseband model of the satellite transmission. Let d_k be a modulation symbol and $\mathbf{d} = (d_0, \dots, d_{N-1})^T$ the sequence to be sent. The lowpass pulse-shaping filter is defined by its impulse response h_{Tx} . Given the input \mathbf{d} , its output is the signal $x(t)$ which is transmitted on the uplink from the gateway to the satellite. The satellite is also considered non-regenerative, it only performs filtering and amplification.

Since the transmitter is usually equipped with a large and highly directional antenna which ensures a high [Signal-to-Noise Ratio \(SNR\)](#) at the satellite transponder input, the uplink is considered noiseless.

Let $y(t)$ be the transponder output to be transmitted on the downlink. We model the nonlinear characteristic of the transponder by \mathcal{R} . As a consequence, $y(t) = \mathcal{R}[x(t)]$.

We assume a *single-carrier-per-transponder* rather than multiple-carriers-per-transponder. Its advantage is to operate closer to the saturation which improves the amplifier efficiency [19]. In that case, each transponder is allocated a frequency bandwidth which does not overlap with adjacent subbands allocated to other transponders.

However, due to channel impairment, signals transmitted by other transponders operating on adjacent bandwidth generate interference called [Adjacent Carrier Interference \(ACI\)](#) which superposes to the signal under concern.

As a consequence, the signal received by the end-user denoted by $r(t)$ reads:

$$r(t) = y(t) + y_{ACI}(t) + b(t), \quad (2.1)$$

where $b(t)$ is the [Additive White Gaussian Noise \(AWGN\)](#) and in the simulations, the interference $y_{ACI}(t)$ is generated as a time-delayed and frequency-shifted version of $y(t)$ that is to say:

$$y_{ACI}(t) = \sum_{\substack{m=-M_{AC}/2 \\ m \neq 0}}^{M_{AC}/2} y(t - \tau_m) e^{2j\pi m \Delta_f t} \quad (2.2)$$

with M_{AC} the number of interfering signals from adjacent transponders (even), Δ_f the frequency offset between subcarriers, and τ_m the time-delay for the m -th

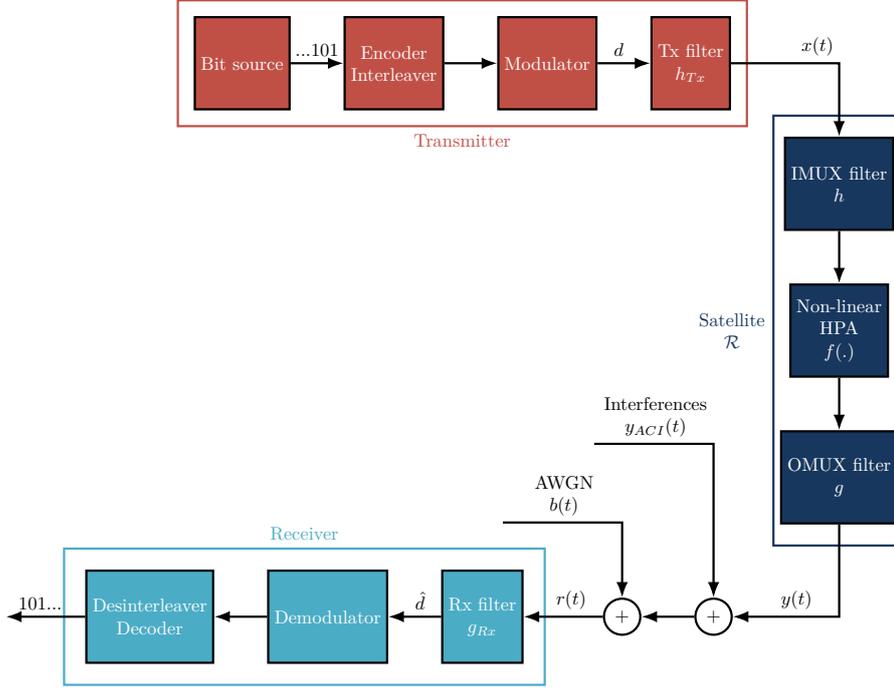


Figure 2.1: Communication Chain

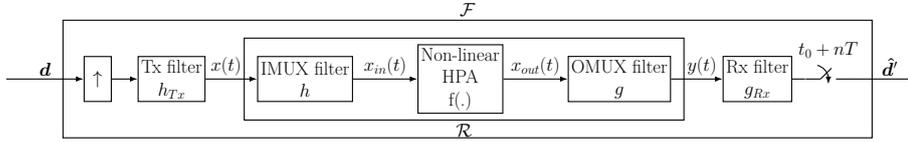


Figure 2.2: Illustration of the transmission model at signal and symbol rate

carrier. The delays τ_m have to be chosen such that the adjacent carriers are not synchronized with the receiver so the carriers might be considered uncorrelated.

At the receiver r is filtered by a low-pass filter whose impulse response is denoted by g_{Rx} . Let r_f stand for the received filter output. r_f is sampled and passed through a decision-making device to produce the symbol estimate sequence $\hat{\mathbf{d}}$. We denote by \mathcal{F} the nonlinear characteristic of the equivalent device which would produce $\hat{\mathbf{d}}$, the sequence of symbols, from \mathbf{d} .

Let us mention that in practice a return link from the satellite to the gateway is used to identify \mathcal{R} or \mathcal{F} and adapt the predistortion.

Figure 2.1 and 2.2 illustrates respectively, the communication chain and the transmission model at both rates.

The communication chain corresponds to the DVB-Sx reference simulation scenario for DTH broadcasting described in [20] without the ACI on the uplink.

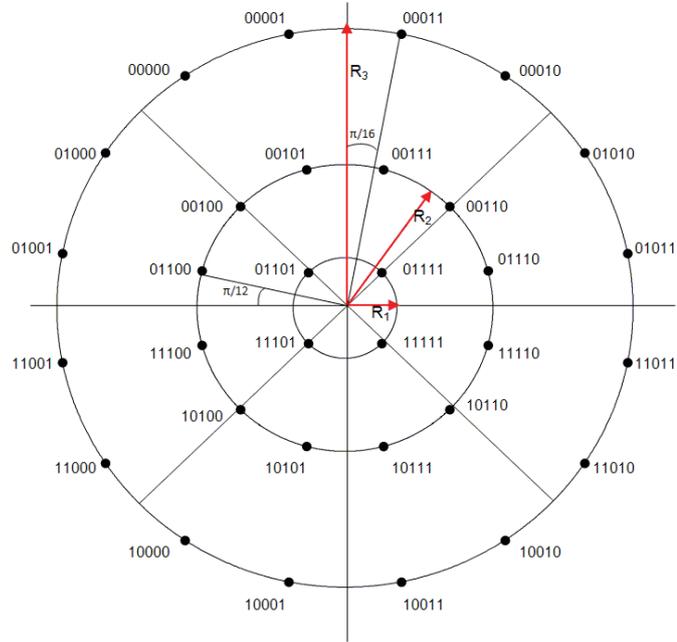


Figure 2.3: 32-APSK [11]. ©European Telecommunications Standard Institute 2014. Further use, modification, copy and/or distribution are strictly prohibited.

2.3 Specificities of the DVB-S2(X) and beyond

In this section, the specificities of the physical layer of the DVB-S2(X) are presented, notably the constellation used, the pulse shaping filter and also the SC-OFDM technique which is a potential candidate for future standards.

2.3.1 Constellation

The DVB-S only considers the QPSK. This modulation has a constant envelope and therefore is more resilient toward nonlinearities. In the next iteration of the standard, higher order modulations are proposed. Instead of the commonly used Quadrature amplitude modulation (QAM), APSK constellations are favoured since APSK constellations tend to have lower envelope power [21]. In the case of APSK constellation, the symbols are mapped on multiple concentric rings.

In the following, the 4+12-APSK and the 4+12+16-APSK will be referred to as the 16-APSK and 32-APSK. The ratios between rings correspond to the highest Modulation/coding spectral efficiency (or equivalently the lowest coding rate) in [11], [12].

Figure 2.3 illustrates the 32-APSK.

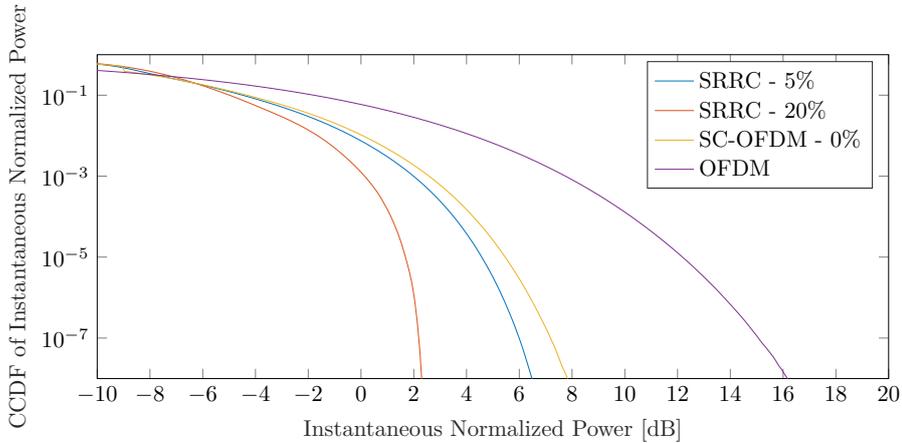


Figure 2.4: Complementary cumulative distribution function of the Instantaneous Normalized Power for different pulse shaping filters - QPSK

2.3.2 Pulse shaping filter

Squared Root Raised Cosine

The [Square Root Raised Cosine \(SRRC\)](#) is a commonly used pulse shaping filter in telecommunications. If both transmitter and receiver filters are SRRC filters, then the convolution $h_{Tx} * g_{Rx}$ verifies the Nyquist criterion.

The frequency response of the SRRC filter denoted by $H_{SRRC}(f)$ is given by [10]:

$$H_{SRRC}(f) = \begin{cases} 1 & \text{for } |f| < \frac{1-\beta}{2T_s} \\ \sqrt{\frac{1}{2} + \frac{1}{2} \sin\left[\frac{\pi T_s}{\beta} \left(\frac{1}{2T_s} - |f|\right)\right]} & \text{for } \frac{1-\beta}{2T_s} \leq |f| \leq \frac{1+\beta}{2T_s} \\ 0 & \text{for } |f| \geq \frac{1+\beta}{2T_s} \end{cases}, \quad (2.3)$$

where β is the roll-off and T_s is the symbol period.

The sharper the roll-off, the bigger the [PAPR](#) [22].

In the [DVB-S2X](#), the roll-off can be as sharp as $\beta = 5\%$, which makes the transmission more vulnerable to nonlinearities.

2.3.3 SC-OFDM

Satellite communications are expected to be a part of the [5G](#) standardized by the [Third Generation Partnership Project \(3GPP\)](#). The [Single-Carrier Frequency Division Multiple Access \(SC-FDMA\)](#), which is used on the uplink of the fourth generation ([Long Term Evolution \(LTE\)](#)) of [Universal Mobile Telecommunications System \(UMTS\)](#), since 2009 [23]. Its broadcast version, the [Single Carrier Orthogonal Frequency Division Multiplexing \(SC-OFDM\)](#) is compatible with a 0% roll-off, which would enhance the spectral efficiency, and therefore it can be seen as a potential technique for future standards [24], [25].

Moreover, the [SC-OFDM](#) has a low single-carrier-like [PAPR](#) and is more resilient to phase noise [24]. Likewise, Figure 2.4 shows that the instantaneous

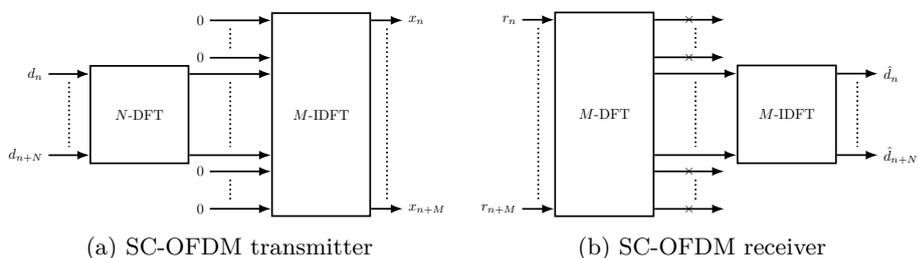


Figure 2.5: SC-OFDM transmitter and receiver

normalized power of the SC-OFDM is lower than one of the Orthogonal Frequency Division Multiplexing (OFDM).

The SC-OFDM is often presented as a Discrete Fourier Transform (DFT)-precoded Orthogonal Frequency Division Multiple Access (OFDMA) which can be confusing since the SC-OFDM is a *single carrier* modulation.

As a consequence, its instantaneous power distribution doesn't follow a Gaussian distribution [26]. The expression of the instantaneous power distribution for SC-OFDM is derived in [27].

An N -DFT is applied to the symbols d_n , subsequently, zero-padding is done on both sides. Then, an M -IDFT is carried out. The number of zeros inserted depends on the oversampling factor. The oversampling factor is defined by the ratio $M/N (\geq 1)$. At the receiver, the reciprocal process is carried out. First, an M -DFT is done on the oversampled signal, then, the N central "subcarriers" of interest are selected and an N -IDFT is performed, yielding the estimate \hat{d} .

Figure 2.5a and 2.5b illustrate, respectively, the SC-OFDM transmitter and receiver.

2.4 Description of the transponder model

The satellite transponder is composed of three elements: the Input MultipleXer (IMUX), the non-linear high power amplifier (HPA) and the Output MultipleXer (OMUX).

2.4.1 IMUX and OMUX filters

The IMUX filter has to extract the uplink carrier of interest and remove interference generated by the adjacent channels. This filter will be referred as h in the following.

The discrete impulse response of IMUX filter is described by the vector $\mathbf{h} = (h_{-L}, \dots, h_0, \dots, h_L)^T$ of length $(2L + 1)$ and is associated to the $M \times M$

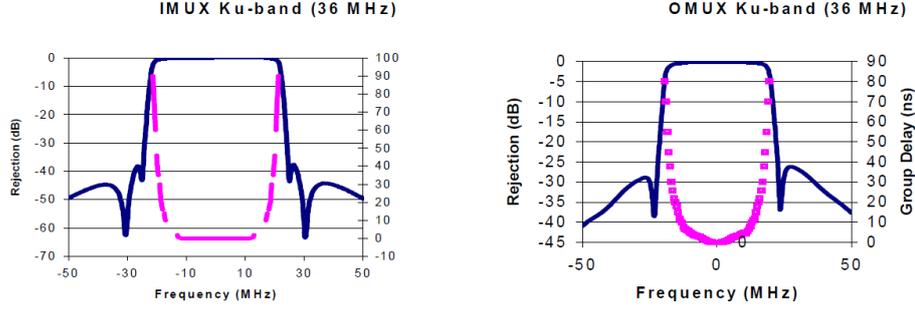


Figure 2.6: IMUX and OMUX characteristics [11]. ©European Telecommunications Standard Institute 2014. Further use, modification, copy and/or distribution are strictly prohibited.

Toeplitz matrix \mathbf{H} defined by:

$$\mathbf{H} = \begin{pmatrix} h_0 & h_{-1} & \cdots & h_{-L} & 0 & \cdots & \cdots & \cdots & 0 \\ \vdots & \ddots & \vdots \\ h_L & \ddots & 0 \\ 0 & \ddots & h_{-L} \\ \vdots & \ddots & \vdots \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \cdots & \cdots & 0 & h_L & \cdots & h_1 & h_0 \end{pmatrix} \quad (2.4)$$

The **OMUX** filter aims to reduce the interference between carriers on the downlink channel by filtering the spectral regrowth generated by the non-linear amplification.

The discrete impulse response of OMUX filter is described by the vector $\mathbf{g} = (g_{-P}, \dots, g_0, \dots, g_P)^T$ of length $(2P + 1)$ and is associated to the $M \times M$ Toeplitz matrix \mathbf{G} .

$$\mathbf{G} = \begin{pmatrix} g_0 & g_{-1} & \cdots & g_{-P} & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ g_P & \ddots & \ddots & \ddots & \ddots & \ddots & 0 \\ 0 & \ddots & \ddots & \ddots & \ddots & \ddots & g_{-P} \\ \vdots & 0 & \ddots & \ddots & \ddots & \ddots & \ddots \\ \vdots & \vdots & \ddots & \ddots & \ddots & \ddots & \ddots \\ 0 & 0 & 0 & 0 & 0 & g_1 & g_0 \end{pmatrix} \quad (2.5)$$

Figure 2.6 illustrated the **IMUX** and **OMUX** characteristics.

2.4.2 Amplifiers

Let us now focus on the on-board power amplifier. It amplifies the received signal in order to cope with the path loss over the downlink.

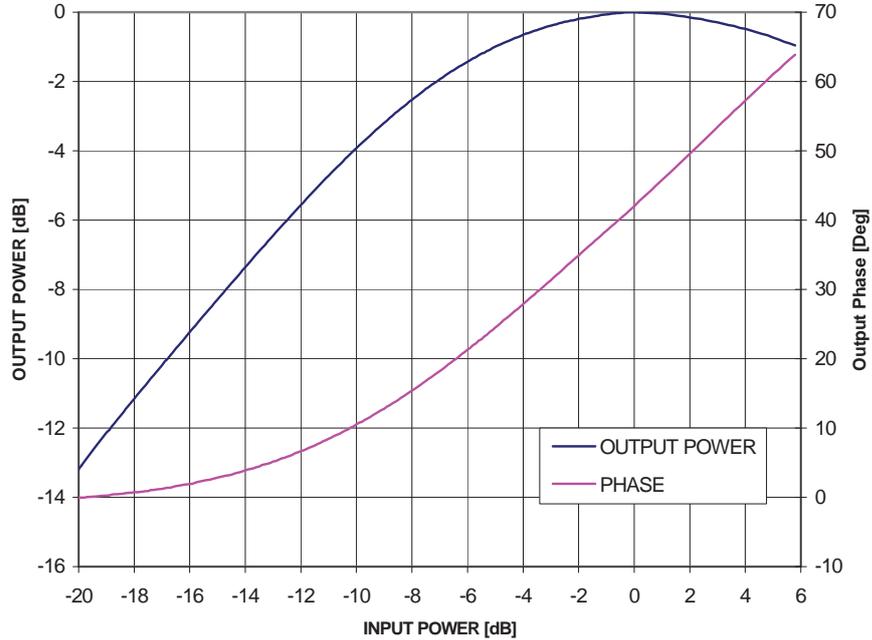


Figure 2.7: AM-AM and AM-PM characteristics proposed in the DVB-S2 standard [11]. ©European Telecommunications Standard Institute 2014. Further use, modification, copy and/or distribution are strictly prohibited.

For satellite communication applications, we focus on two types of amplifiers: **Travelling-Wave Tube Amplifier (TWTA)** and **Solid State Power Amplifier (SSPA)**. **TWTA** are mainly used over the C , K_u and K_a band whereas the **SSPA** are used over the L and C band [13]. Both amplifiers are considered to be memoryless and have frequency-independent characteristics.

Given an input signal $x_{in}(t) = \rho(t)e^{j\theta(t)}$, the complex output signal $x_{out}(t)$ of the amplifier is modelled as follows:

$$x_{out}(t) = f(x_{in}(t)) \quad (2.6)$$

$$= A(\rho(t))e^{j(\Omega(\rho(t))+\theta(t))}. \quad (2.7)$$

where $f(\cdot)$ is the nonlinear function, $\theta(t)$ is the instantaneous input phase and $\rho(t)$ is the instantaneous input amplitude with $A(\cdot)$ and $\Omega(\cdot)$ being respectively the **AM-AM** and **AM-PM** modulation.

Figure 2.7 and 2.8 illustrate both the **AM-AM** and **AM-PM** for respectively, the **TWTA** as defined in the **DVB-S2** and for multiple Saleh model's characteristics.

For ease of notation, $f(\cdot)$ can either be applied to a scalar or to a vector (e.g. given $\mathbf{Z} = [z_0 \cdots z_{M-1}]^T$, the notation $f(\mathbf{Z})$ stands for a vector of length M , whose i -th component equals $f(z_i)$).

TWTA **TWTA** tends to be less linear than **SSPA**. However, its **Direct Current (DC) to Radio Frequency (RF)** conversion efficiency and maximum output power in terms of watts keep it relevant in most applications [13].

The DVB-S2 gives the characteristics of a non-linearized **TWTA** [11, p. 70].

A commonly used model for **TWTA** is described in [28]. The Saleh model uses four parameters $(\alpha_a, \beta_a, \alpha_\phi, \beta_\phi)$:

$$A(\rho(t)) = \alpha_a \frac{\rho(t)}{1 + \beta_a \rho^2(t)}, \quad (2.8)$$

$$\Omega(\rho(t)) = \alpha_\phi \frac{\rho^2(t)}{1 + \beta_\phi \rho^2(t)}, \quad (2.9)$$

with $\rho^2(t)$ the instantaneous signal power. This model can be made frequency-dependent by adding a filter [28], [29].

SSPA **SSPA** is more linear than **TWTA** and therefore is considered to be less critical than **TWTA**. **SSPA** is often modelled with a Rapp Model. The Rapp model only describes the amplitude distortion [29]. It is mainly used for the **SSPA**.

The amplitude modulation is given by [29]:

$$A(\rho(t)) = \frac{\rho(t)}{\left(1 + \left(\frac{\rho(t)}{\rho_{sat}}\right)^{2P}\right)^{\frac{1}{2P}}}, \quad (2.10)$$

where ρ_{sat} is the saturation amplitude and P is the smoothness factor.

An extended version of this model can be found in [30]. A more general model, the Ghorbani model which is similar to the Saleh model with additional parameters is also used for **SSPA**.

IBO - OBO The **IBO** measures the difference between the mean input power $P_{in,mean}$ and the maximum input power $P_{in,sat}$.

The **IBO** is expressed as:

$$IBO = P_{in,sat} - P_{in,mean} \quad [\text{dB}]. \quad (2.11)$$

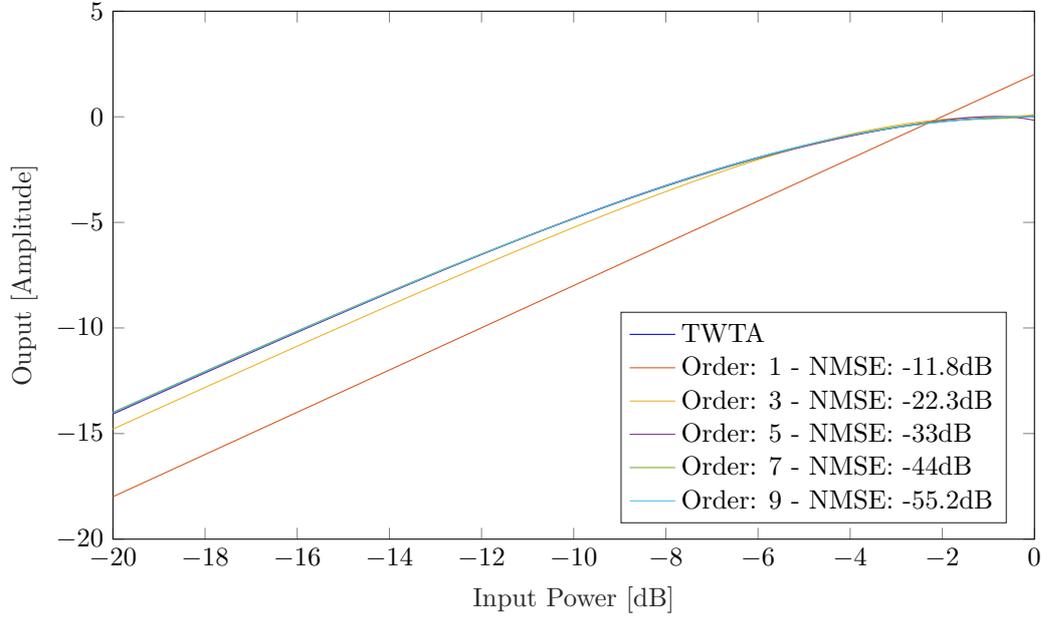
The **OBO** measures the difference between the mean output power $P_{out,mean}$ and the maximum output power of the amplifier $P_{out,sat}$. The **OBO** is expressed as:

$$OBO = P_{out,sat} - P_{out,mean} \quad [\text{dB}]. \quad (2.12)$$

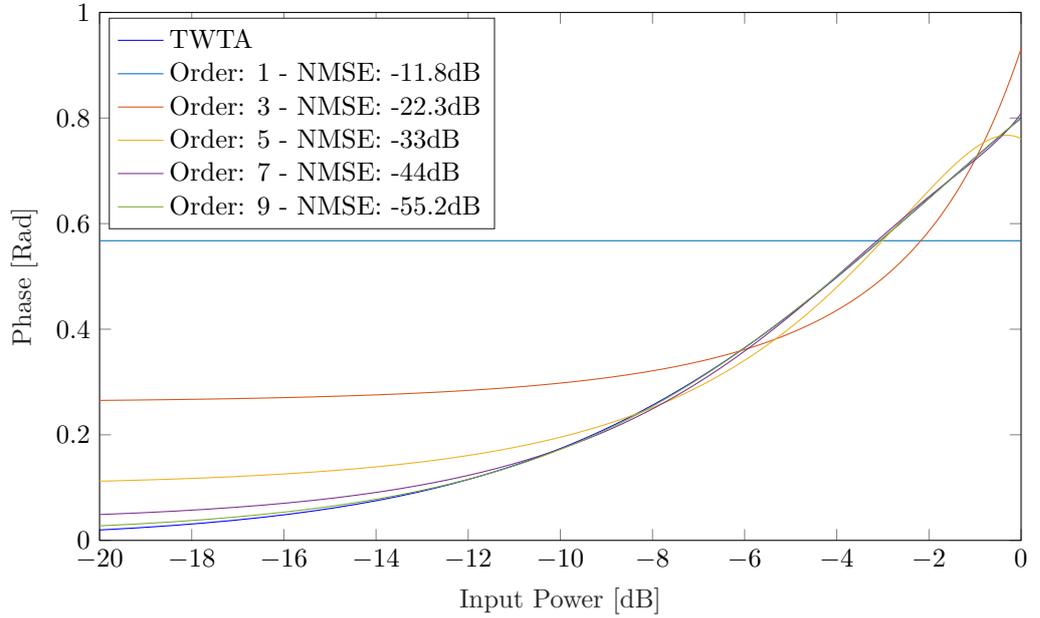
Since $P_{out,sat}$ is fixed by the type of amplifier and the power consumption of the amplifier, the **OBO** can be linked to the power at the receiver. For the same loss on the downlink, the smaller the **OBO**, the more powerful the received signal.

The **OBO** depends on the HPA characteristics and the distribution of the instantaneous power of the input signal. The latter can be linked to the chosen constellation and pulse shaping filter. As a consequence of the Jensen's inequality, $P_{out,mean} \leq f(P_{in,mean})$.

Figure 2.9 illustrates both the **IBO** and the **OBO** with the Saleh model (2,1,2,1.5).



(a) Amplitude characteristics of the Saleh model and its polynomial approximation



(b) Phase characteristics of the Saleh model and its polynomial approximation

Figure 2.8: AM-AM and AM-PM of the Saleh model (2,1,2,1.5) and its polynomial approximation

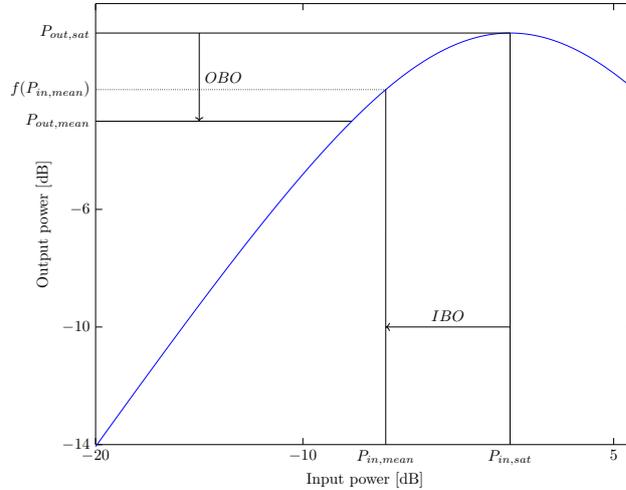


Figure 2.9: Illustration of IBO and OBO over the Saleh model (2,1,2,1.5)

2.5 Modelling of the satellite transponder as a nonlinear system with memory

The satellite transponder can be modelled as a nonlinear filter with memory [31]. In this section, we describe different nonlinear models with memory that can either be used for identification or predistortion. In practice, the satellite transponder has to be identified since the HPA characteristics may vary over time due to notably ageing and temperature changes [13], [21], [32]. We only describe the nonlinear models considered in this thesis. Interested readers can refer to [33] which provides a deep overview of the state-of-the-art of nonlinear models.

2.5.1 Wiener-Hammerstein

The Wiener-Hammerstein (also called LNL for Linear Nonlinear Linear) model is a block-oriented nonlinear model. Figure 2.10 illustrates the scheme. The Wiener-Hammerstein model directly applies to the satellite transponder since it matches its description.

The identification of this type of model requires additional assumptions in order to raise ambiguity linked to the structure [34], [35].

A matrix version of the Wiener-Hammerstein might be used in order to describe \mathcal{R} :

$$\mathcal{R}\{\mathbf{X}_n\} = \mathbf{G}f(\mathbf{H}\mathbf{X}_n), \quad (2.13)$$

where \mathbf{X}_n is the sampled signal vector defined by:

$$\mathbf{X}_n = (x_{n-(L+P)}, \dots, x_n, \dots, x_{n+(L+P)})^T, \quad (2.14)$$

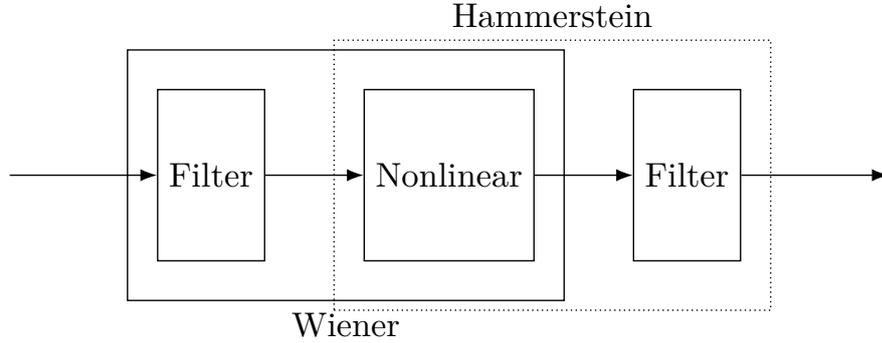


Figure 2.10: Wiener-Hammerstein model

2.5.2 Volterra model

A more commonly used model in the literature since it can be viewed as a Taylor expansion with memory and, therefore, as an extension of the linear case, is the Volterra model [36].

The satellite transponder is assumed to be bandpass and the kernels of the Volterra model are symmetric functions of their arguments, then one can only take into account the odd terms of the Volterra model (and Taylor series for the amplifier characteristics) [31].

This assumption of symmetry doesn't imply any loss of generality [31], [37].

Let us assume that the amplifier characteristics can be expanded as a Taylor series. Then, the satellite transponder can be described by a Volterra model [31], [38].

The nonlinear function of the amplifier is decomposed as follows:

$$f(x_{in}(t)) = \sum_{m=0}^{\infty} \gamma_{2m+1} x_{in}(t) |x_{in}(t)|^{2m}, \quad (2.15)$$

with $x_{in}(t)$ the signal entering the amplifier and γ_{2m+1} the $2m+1$ -th coefficient of the Taylor series.

The input-output relationship for a *Volterra* system can be expressed as [38]:

$$y(t) = \sum_{m=1}^{\infty} \mathcal{Q}_{(m)}[x(t)] \quad (2.16)$$

where

$$\mathcal{Q}_{(m)}[x(t)] \triangleq \underbrace{\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty}}_{2m+1 \text{ integrals}} q_{(2m+1)}(\tau_1, \dots, \tau_{2m+1}) \prod_{i=1}^{m+1} x(t - \tau_i) d\tau_i \cdot \prod_{j=m+2}^{2m+1} x^*(t - \tau_j) d\tau_j \quad (2.17)$$

and $q_{(k)}(\tau_1, \dots, \tau_{2m+1})$ is the k -th order low-pass Volterra kernel.

Those kernels can be expressed as follows:

$$q_{(2m+1)}(\tau_1, \dots, \tau_{2m+1}) = \gamma_{2m+1} \int_{-\infty}^{\infty} g(\tau) \prod_{k=1}^{m+1} h(\tau_k - \tau) \cdot \prod_{s=m+2}^{2m+1} h^*(\tau_s - \tau) d\tau \quad (2.18)$$

where γ_{2m+1} is the $2m + 1$ -th coefficient of the Taylor expansion of the nonlinearity.

The length of the equivalent filter increases exponentially with the order k since the **IMUX** filter is convoluted with itself k -times. Therefore, the shorter the impulse response h , the less kernels to be taken into account when sampled. However, even if the equivalent impulse response is infinite, the output only depends of the $(L + P)$ previous and following samples at any given instant.

In practice, the Volterra model is sampled and truncated. The Volterra model can be expressed as:

$$y_n = \sum_{m=0}^{K'} \underbrace{\sum_{n_1=-(L+P)}^{L+P} \cdots \sum_{n_{2m+1}=-(L+P)}^{L+P}}_{2m+1 \text{ sums}} q_{(2m+1), n_1, \dots, n_{2m+1}} \prod_{k=1}^{m+1} x_{n-n_k} \prod_{s=m+2}^{2m+1} x_{n-n_s}^* \quad (2.19)$$

where y_n is the n -th output sample, x_{n-n_i} is the $(n - n_i)$ -th input sample, $q_{(2m+1), n_1, \dots, n_{2m+1}}$ is the $2m+1$ -th order sampled kernel for the lags n_1, \dots, n_{2m+1} , $K' = \frac{K-1}{2}$ depends on the maximum order K , L is the memory length of the **IMUX** filter and P is the memory length of the **OMUX** filter.

The $2m + 1$ -th sampled kernel $q_{(2m+1), n_1, \dots, n_{2m+1}}$ reads:

$$q_{(2m+1), n_1, \dots, n_{2m+1}} = \gamma_{2m+1} \sum_{l=-\infty}^{\infty} g_l \prod_{k=1}^{m+1} h_{n_k-l} \prod_{s=m+2}^{2m+1} h_{n_s-l}^* \quad (2.20)$$

Replacing equation (2.20), h by $h_{pre} = h * h_{Tx}$ and g by $g_{post} = g * g_{Rx}$, we obtain, the kernel of \mathcal{F} , which we denote by $\tilde{q}_{(2m+1), n_1, \dots, n_{2m+1}}$ that links the sequences of transmitted and received symbols. Then, equation (2.19) becomes:

$$\hat{d}_n = \sum_{m=0}^{K'} \underbrace{\sum_{n_1=-L'}^{L'} \cdots \sum_{n_{2m+1}=-L'}^{L'}}_{2m+1 \text{ sums}} \tilde{q}_{(2m+1), n_1, \dots, n_{2m+1}} \prod_{k=1}^{m+1} d_{n-n_k} \prod_{s=m+2}^{2m+1} d_{n-n_s}^* \quad (2.21)$$

Based on eq.(2.19) and the fact that only a finite number of samples is needed to compute the output, the Volterra model can be seen as an extension of the linear case since the Volterra model can be written as a matrix operation [36]:

$$y_n = \mathbf{Q}^T \Phi[\mathbf{X}_n] \quad (2.22)$$

with y_n the output sample of the Volterra model and, \mathbf{Q} the kernel vector

defined as follows:

$$\mathbf{Q} = \begin{pmatrix} q(1, -(L+P)) \\ \vdots \\ q(1, (L+P)) \\ q(3, -(L+P), -(L+P), -(L+P)) \\ \vdots \\ q(3, L+P, L+P, L+P) \\ \vdots \\ q(K, L+P, \dots, L+P) \end{pmatrix}, \quad (2.23)$$

$\Phi(\cdot)$ is an operator which computes the nonlinear combination of the input vector \mathbf{X}_n . $\Phi[\mathbf{X}_n]$ is defined as:

$$\Phi[\mathbf{X}_n] = \begin{pmatrix} x_{n-(L+P)} \\ \vdots \\ x_n \\ \vdots \\ x_{n+L+P} \\ x_{n-(L+P)} |x_{n-(L+P)}|^2 \\ x_{n-(L+P)} x_{n-(L+P)} x_{n-(L+P)}^* \\ \vdots \\ x_{n+(L+P)} |x_{n+(L+P)}|^2 \\ \vdots \\ x_{n+(L+P)} |x_{n+(L+P)}|^{K-1} \end{pmatrix}. \quad (2.24)$$

The number of kernels $C_{Volterra}(K, L+P)$ (which is related to the complexity of the model) is given by:

$$C_{Volterra}(K, L+P) = \sum_{m=0}^{K'} \binom{2(L+P)+1+m}{m+1} \binom{2(L+P)+m}{m} \quad (2.25)$$

where $\binom{n}{k}$ is the binomial coefficient. The number of kernels for odd and even orders in the real case has been derived in [39].

2.5.3 Memory Polynomial

The Memory Polynomial (MP) has the same mathematical expression as the Hammerstein model, if the nonlinearity is decomposed as in 2.15, except that it's doesn't handle a block model. It can also be seen as a special case of the previously presented Volterra model where only the diagonal terms are taken [40] and it can be described by using a matrix formulation identical to the one used in eq.(2.22). Its input-output relationship can be expressed as:

$$y(t) = \sum_{\substack{k=1 \\ \text{odd}}}^K \sum_{l=-(L+P)}^{L+P} a_{k,l} x(t) |x(t-l)|^{k-1}. \quad (2.26)$$

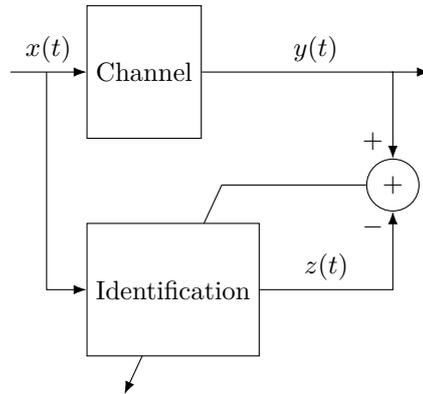


Figure 2.11: Identification principle

The number of kernels $C_{MP}(K, L)$ (which is related to the overall complexity) is proportional to the order used for the nonlinearity and to the memory depth:

$$C_{MP}(K, L) = \frac{1}{2}(2L + 1)(K + 1). \quad (2.27)$$

2.5.4 Identification

The identification of the channel characteristics is necessary since the components, especially the **High Power Amplifier (HPA)** may vary over time due to, for instance, ageing and temperature change [21], [32]. Moreover, two **HPA** with the same design may differ in their characteristics due to difference in the making [13].

The identification for the Volterra model and the Memory polynomial model is performed in a similar way as for the linear case. The identification for those models aims to estimate the kernel vector \mathbf{Q} (defined in eq.(2.23)). Figure 2.11 illustrates the identification process. The identification can be performed with a **LMS** or **Recursive Least Squares (RLS)** method. Those methods are presented in a comprehensive way in [36].

In some cases, Volterra kernels may be degenerated resulting in non-invertible correlation matrix ($\mathbf{R}_\Phi = \Phi[\mathbf{X}_n]^\dagger \Phi[\mathbf{X}_n]$) which makes the identification impossible to perform.

Figure 2.8 shows the polynomial approximation of a Saleh model. As the order increases, the approximation becomes better.

2.6 Impairments due to the satellite transponder

In Section 2.5, the satellite transponder has been described as a nonlinear system with memory involving **AM-AM** and **AM-PM**.

As a consequence, the received signal will be degraded due to three main effects [41]:

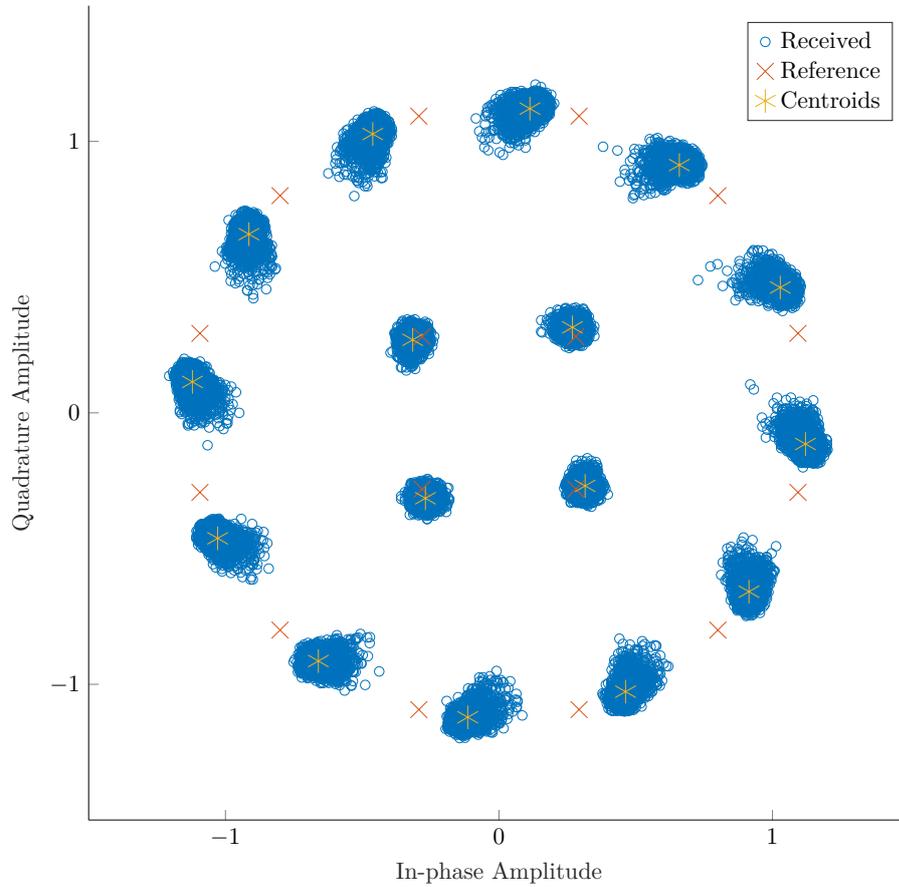


Figure 2.12: Clustering and wrapping for 16-APSK – Pulse shaping: $\beta = 0.05$ – IBO = 12dB - Noiseless without IMUX and OMUX filters

- *Wrapping* that is to say the displacement of the constellation centroids (with respect to the transmitted constellation) due to [AM-AM](#) and [AM-PM](#).
- *Clustering* which corresponds to the spreading of the received symbols due to the mismatch between the pulse shaping filter and the matched receiver filter since the amplification is non-linear.
- *Spectral regrowth* which increases the bandwidth occupied by the signal and might lead to spillage over adjacent carriers.

Figure 2.12 shows the wrapping effect, the centroids (green) of the received clusters are separated from the reference points of the constellation (red), and the clustering effect (blue). It can be noticed that the distance between the radii is reduced and the displacement, phase shift and the clustering effect depend on the radius.

In the following, we demonstrate the statistical properties of the wrapping and the clustering effect for a third order Volterra model and we show the impact of smaller IBO and nonlinearities on the spectral regrowth through simulation.

2.6.1 Wrapping

Let us illustrate the wrapping effect from the equivalent discrete baseband model (equation (2.21)).

We assume zero-mean independent and identically distributed modulation symbols.

$$\begin{aligned}
 \mathbb{E}[\hat{d}_n | d_n = d] &= \tilde{q}_{(1),0} d \\
 &+ \tilde{q}_{(3),0,0,0} |d|^2 d \\
 &+ \sum_{\substack{i=-L \\ i \neq 0}}^L \tilde{q}_{(3),i,i,0} R_2^{2,0} d^* \\
 &+ \sum_{\substack{i=-L \\ i \neq 0}}^L (\tilde{q}_{(3),i,0,i} + \tilde{q}_{(3),0,i,i}) \sigma_c d
 \end{aligned} \tag{2.28}$$

where σ_c is the variance of the constellation and $R_2^{2,0}$ is the pseudo-variance of the constellation. The variance and pseudo-variance displaces the centroid due to the clustering effect. It can be noticed that $R_2^{2,0} \neq 0$ only for constellation that has a ring with only two symbols in it.

For higher order model, without the effect of the clustering, the sum of the central kernels ($\tilde{q}_{(2m+1),0,\dots,0}$) appear and using equation (2.15), we have:

$$\sum_{m=0}^{\infty} \tilde{q}_{(2m+1),0,\dots,0} d |d|^{2m} = \sum_{m=0}^{\infty} \gamma_{2m+1} g_{Post,0} h_{Pre,0} |h_{Pre,0}|^{2m} d |d|^{2m} \tag{2.29}$$

$$= g_{Post,0} f(h_{Pre,0} d) \tag{2.30}$$

where $h_{Pre,0}$ and $g_{Post,0}$ are the central tap of respectively of the impulse response before and after the amplifier, h_{Pre} and g_{Post} . The characteristics of the amplifier depends on the amplitude of the input signal, then for a multiple energy rings constellation, all the centroids are not displaced by the same amount.

2.6.2 Clustering

Let us now describe the clustering effect. The clustering effect can be seen as an additional noise due to the distortion. It can also be important to study the clustering effect in order to define the decision boundaries, which is important for the computation of [Log-Likelihood Ration \(LLR\)](#).

As we work with complex constellation, we need to derive both the variance and the pseudo-variance in order to characterize the clustering effect. The reader interested in the full derivation can report to appendix A. The derivation is valid for a large number of constellations that can be written as:

$$\{ \{ a_i e^{2j\pi \frac{k}{n_i}}, i = [1, R], a_i \in \mathbb{C}, \forall n_i \in \mathbb{N}^* \setminus \{1\}, k = [0, n_i - 1] \}, \forall R \in \mathbb{N}^* \} \tag{2.31}$$

where R is the number of rings, n_i (≥ 2) is the number of symbols of the i -th ring, $a_i \in \mathbb{C}$ is the amplitude and phase shift of the i -th ring.

The conditional variance for a third order Volterra series can be expressed as:

$$\begin{aligned} \mathbb{E}[|\hat{d}_n - \mathbb{E}[\hat{d}_n|d_n = d]|^2|d_n = d] &= \mathbf{q}_1\sigma_c^2 + \mathbf{q}_2\sigma_c^2|d|^2 + \Re(\mathbf{q}_3R_2^{2,0}(d^*)^2) + \mathbf{q}_4R_4^{2,2} \\ &+ \mathbf{q}_5\sigma_c^4 + \mathbf{q}_6|R_2^{2,0}|^2 + \mathbf{q}_7\sigma_c^2|d|^4 \\ &+ \Re(\mathbf{q}_8R_2^{2,0}d^2)|d|^2 + \mathbf{q}_9\sigma_c^4|d|^2 \\ &+ \mathbf{q}_{10}|R_2^{2,0}|^2|d|^2 + \mathbf{q}_{11}R_6^{3,3} + \mathbf{q}_{12}\sigma_c^6 \\ &+ \mathbf{q}_{13}|R_2^{2,0}|^2\sigma_c^2 + \mathbf{q}_{14}R_4^{2,2}\sigma_c^2 \end{aligned} \quad (2.32)$$

with σ_c^2 being the variance of the constellation, and $R_m^{k,l} = \mathbb{E}[d^k(d^*)^l]$ with $(m = k + l)$ being a high order statistics and \mathbf{q} is a complex coefficient resulting of the sum of product of the first and third order kernel \tilde{q} .

We see that the conditional variance of the received symbol depends on the statistics of the chosen constellation ($\sigma_c^2, R_2^{2,0}, R_4^{2,2}, R_6^{3,3}$), the modulus of the symbol under consideration ($|d|$) and the phase of the considered symbol. The symbols on the outer rings of multiple rings constellation tend to be more affected by the distortion than the one on the inner rings since the modulus $|d|$ is greater on the outer rings. This effect can be seen in Figure 2.12. Even if in that case, the nonlinear is only amplitude-dependent, the conditional variance of the system with memory is both dependent on the amplitude and the phase. That is to say that the relative position of the symbol under interest with rings containing only two symbols is affecting the conditional variance.

To illustrate this phenomenon, we simulate an 2+4+2 APSK, which is composed of what will refer as an inner **Binary Phase Shift Keying (BPSK)** (labelled 1 and 2 in Figure 2.14) an evenly distributed **Quadrature Phase Shift Keying (QPSK)** (labelled 3,4,5 and 6 in Figure 2.14) and an outer **BPSK** (labelled 7 and 8 in Figure 2.14). This constellation is similar to the one standardized in the **DVB-S2X**, except that, the middle rings with four symbols are evenly distributed as in a **QPSK**. In this case, the second order statistics $R_2^{2,0} = \frac{1}{4}(a_1^2 + a_3^2)$ which is of similar magnitude of the variance $\sigma_c = \frac{1}{8}(2|a_1|^2 + 4|a_2|^2 + 2|a_3|^2)$. The nonlinear amplifier is a Saleh amplifier approximates to the third order - see (fig.2.8), the transmitter and receiver filters are **SRRC** filter with a 5% roll-off.

Figure 2.13 shows the evolution of the conditional variance of the different symbol of the constellation for different phase offset at the origin of the **QPSK**. In that case, the conditional variance is dependent on the phase and the modulus of the symbols.

Figures 2.14a and 2.14b shows the constellation for two different phase offset (0 and $\frac{\pi}{4}$).

The conditional pseudo-variance reads as:

$$\begin{aligned}
 \mathbb{E}[(\hat{d}_n - \mathbb{E}[\hat{d}_n|d_n = d])^2|d_n = d] &= \mathbf{q}'_1 R_2^{2,0} + \mathbf{q}'_2 R_2^{2,0}|d|^2 + \mathbf{q}'_2 \sigma_c^2 d^2 + \mathbf{q}'_3 \sigma_c^2 R_2^{2,0} \\
 &+ \mathbf{q}'_4 R_2^{2,0}|d|^4 + \mathbf{q}'_5 R_2^{2,0}|d|^2 d^2 \\
 &+ \mathbf{q}'_6 R_2^{0,2} d^4 + \mathbf{q}'_7 R_4^{2,2} d^2 + \mathbf{q}'_8 d^2 \sigma_c^4 \\
 &+ \mathbf{q}'_9 |R_2^{2,0}|^2 d^2 + \mathbf{q}'_{10} R_4^{4,0} (d^*)^2 \\
 &+ \mathbf{q}'_{11} (R_2^{2,0})^2 (d^*)^2 + \mathbf{q}'_{12} R_4^{4,0} R_2^{0,2} \\
 &+ \mathbf{q}'_{13} |R_2^{2,0}|^2 R_2^{2,0} + \mathbf{q}'_{14} R_4^{2,2} R_2^{2,0} \\
 &+ \mathbf{q}'_{15} \sigma_c^4 R_2^{2,0} \tag{2.33}
 \end{aligned}$$

Assuming that the (transmitter, **IMUX,OMUX** and receiver) filters have real coefficients and according to equation (2.20), then, those properties will depend on the product of coefficients γ_{2m+1} of Taylor expansion of the characteristics of the amplifier, d^2 and high orders statistics of the constellation, namely $R_2^{2,0}$ and $R_4^{4,0}$. Therefore, in most cases, it can be expect that the conditional distribution to be elliptic.

The influence of the channel can be seen between Figures 2.14a and 2.14c and between Figures 2.14b and 2.14d where two approximation to the third order of Saleh model are used respectively $(2,1,2,1.5)$ and $(2,1,\frac{\pi}{6},1)$ and the difference in the cluster forms due to the phase offset between Figures 2.14a and 2.14b and, for another Saleh model between Figures 2.14c and 2.14d.

However, if the chosen constellation contains enough symbols per rings, in our case of third order Volterra systems at least six then the high-orders statistics $R_2^{2,0} = 0$ and $R_4^{4,0} = 0$. Therefore, in that case, the conditional distribution is invariant by rotation since:

$$\mathbb{E}[|\hat{d}_n - \mathbb{E}[\hat{d}_n|d_n = de^{j\alpha}]|^2|d_n = de^{j\alpha}] = \mathbb{E}[|\hat{d}_n - \mathbb{E}[\hat{d}_n|d_n = d]|^2|d_n = d] \tag{2.34}$$

$$\mathbb{E}[(\hat{d}_n - \mathbb{E}[\hat{d}_n|d_n = de^{j\alpha}])^2|d_n = de^{j\alpha}] = \mathbb{E}[(\hat{d}_n - \mathbb{E}[\hat{d}_n|d_n = d])^2|d_n = d] e^{2j\alpha} \tag{2.35}$$

for any $\alpha \in \mathbb{R}$. Then, received symbols for two different transmitted symbols on the same rings, generates the same pattern rotated angularly by the angle between the two different transmitted symbols.

As for the whole constellation, the pseudo-variance is non-zero ($\mathbb{E}[\mathbb{E}[(\hat{d}_n - \mathbb{E}[\hat{d}_n|d_n = d])^2|d_n = d]] \neq 0$), therefore, the received symbol is improper and has correlated real and imaginary part and is circular, except if the high orders statistics $R_2^{2,0}$ and $R_4^{4,0}$ are equal to zero. This show the need for using widely linear filter for equalization of received symbols [42].

Similar result can be easily derived for higher order Volterra model based on the previous derivation. In the case of higher order Volterra model, higher order statistics of the transmitted constellation are required.

2.6.3 Spectral regrowth

The spectral regrowth is due to the intermodulation with the signal itself. The spectral regrowth can lead to spillage over adjacent carriers. The spectral regrowth increases with the **IBO**, and therefore the **OBO**. Figure 2.15 shows the average **Power Spectral Density (PSD)** for different **IBO** values. We can see that

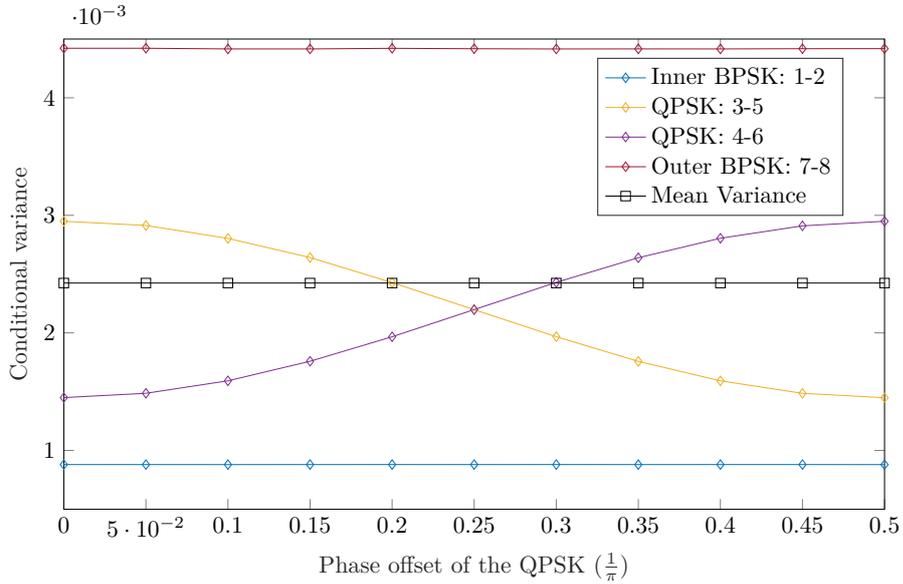


Figure 2.13: Conditional variance for different phase offset of the QPSK.

as the **IBO** decreases, the out-of-band emission increases. Figure 2.16 shows the power spectral density at the output of the amplifier for the same **IBO** as the order of approximation of the amplifier increases. The figure shows that the spectral support widens as the order of nonlinearity increases (only the odd orders are considered). The characteristics of the amplifier approximated up to a certain order can be seen in Figure 2.8a.

2.7 Performance metrics

Three figures of merit are of main interest in order to evaluate the performance of the system:

- **TD** to determine the overall performance of the system,
- **NMSE** to measure the in-band distortion,
- **ACI** factor to estimate the spectral regrowth level.

2.7.1 Total degradation

The **TD**, for a given BER, can be expressed as:

$$TD = OBO + \left. \frac{E_b}{N_0} \right|_{NL} - \left. \frac{E_b}{N_0} \right|_{AWGN} \quad [\text{dB}]. \quad (2.36)$$

It measures the degradation due to the non-linearity and the loss of power owing to the **OBO** defined in Section 2.4.2. $\left. \frac{E_b}{N_0} \right|_{NL}$ and $\left. \frac{E_b}{N_0} \right|_{AWGN}$ represent respectively, the signal-to-interference-plus-noise (SINR) in the nonlinear case and in the AWGN case. Their difference measures the degradation induced by the distortion.

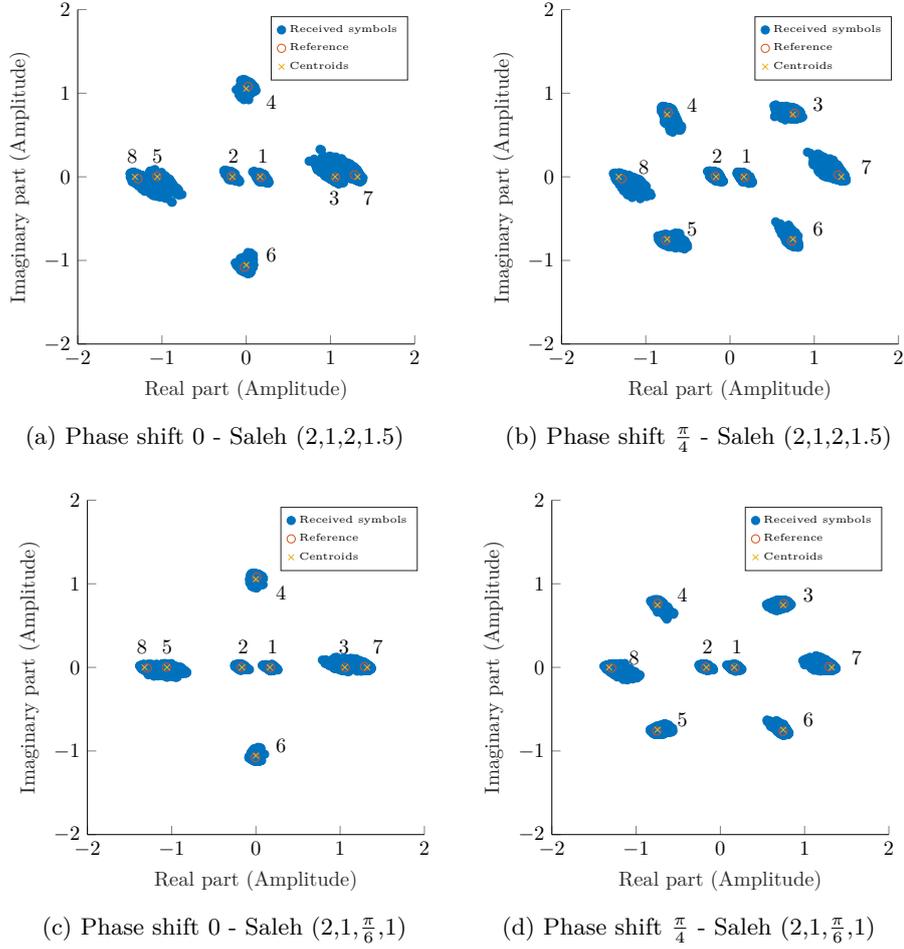


Figure 2.14: 2+4+2 APSK for different phase offset of the QPSK and different Saleh model.

2.7.2 NMSE

The [NMSE](#), quantifies the in-band distortion at symbol rate and it is expressed as:

$$NMSE = 10 \log \mathbb{E} \left(\frac{\sum_{n=0}^N |\hat{d}'_n - d_n|^2}{\sum_{n=0}^N |d_n|^2} \right) \quad [\text{dB}], \quad (2.37)$$

where \hat{d}' is the noiseless and interference-free k -th receiver filter output (*cf* figure 2.2) and d_n is the n -th transmitted symbol, which corresponds to the ideal receiver filter output.

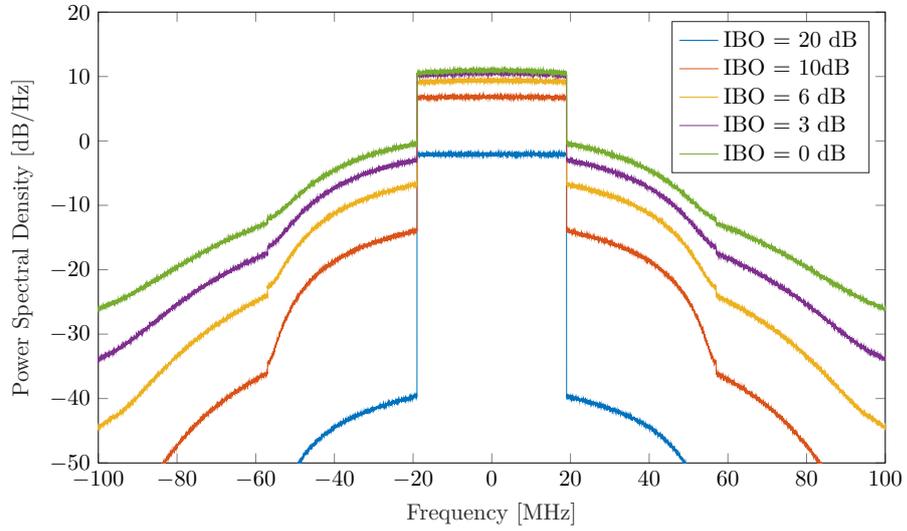


Figure 2.15: Power spectral density over frequency for different IBO - Saleh model (2,1,2,1.5) - SC-OFDM - Symbol rate: 38 MBd.

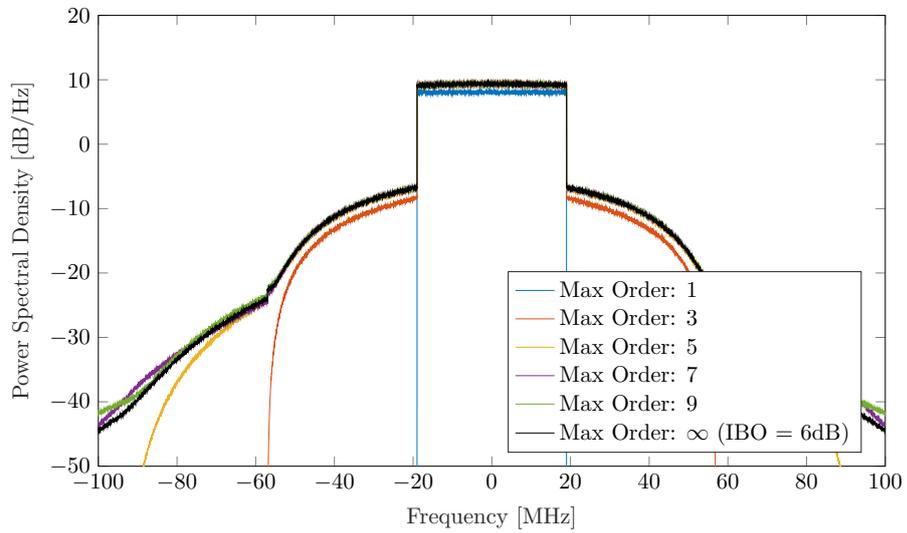


Figure 2.16: Power spectral density over frequency for different order of approximation of the function - Saleh model (2,1,2,1.5) - SC-OFDM - Symbol rate: 38 MBd.

2.7.3 ACI factor

The **ACI** factor is used to measure the interference generated on the considered carrier by the remaining spectral regrowth of adjacent carriers after the signal goes through the receiver's filter. As mentioned in [30], the **ACI** can be expressed as:

$$ACI = 10 \log \mathbb{E} \left(\frac{\sum_{n=0}^{N_s} |(y_{ACI} * h_{RX})_n|^2}{\sum_{n=0}^{N_s} |(y * h_{RX})_n|^2} \right) \quad [\text{dB}], \quad (2.38)$$

where $(a * b)_n$ is the n -th sample of the convolution of a and b and N_s is the number of samples.

The **ACI** is the inverse of the **Signal-to-Interference Ratio (SIR)**.

2.8 Conclusion

This chapter dealt with satellite communications. We described the satellite transmission scenario and introduced the notations that will be used throughout the report. This scheme is similar to the reference DVB-S(X) **DTH** simulation. In this scenario, the receivers are cost-driven, therefore, impairments of the satellite channel should be rather dealt at the transmitter.

Then, we went through the specificities of the **DVB-S2** and **DVB-S2X**, which define the transmitter. Then, following the signal chain, we described the satellite transponder. Multiple amplifiers that might be found in satellite were presented.

Subsequently, multiple ways to model the satellite transponder which behaves as a nonlinear system with memory was described and identification procedure was presented. The identification has to be performed in practice.

Afterwards, based on the mathematical modelling of the transponder, we discussed the three principal impairments generated by the transponder. We analysed the effect of the in-band distortion and showed the link between the order of nonlinearities and the spectral regrowth.

Finally, the performance metric were defined.

Based on the observations made in this chapter, we will interest to compensation of the nonlinearities at the transmitter, namely predistortion.

After this introduction and modelling of satellite communications, we will now take an interest in the compensation of nonlinearities at the transmitter, in other words, in next chapters we will investigate predistortion techniques.

Chapter 3

State-of-the-art of predistortion techniques

“Time is not your friend. It goes no matter what you say. No matter what I think, it won’t wait for me. It won’t go ahead of me. So when I wish it would go faster, it just stays there. And when I wish it would slow down, it won’t listen to me. It makes me ride shotgun. No drive, just ride.”

Josh Homme as himself
American Valhalla. Directed by Andreas
Neumann and Joshua Homme.
Eagle Rock Films. 2017

Contents

3.1 Introduction	48
3.2 Direct predistortion method	48
3.2.1 P-th order inverse	48
3.2.2 Global compensation	49
3.2.3 Learning architecture	50
3.2.4 Neural network as a predistortion technique	53
3.3 Iterative predistortion methods	54
3.3.1 Small variation algorithm	55
3.3.2 Iterative predistortion based on the fixed-point theorem	56
3.3.3 Successive approximation	57
3.4 Conclusion	59

3.1 Introduction

In the previous chapter, we described the impairments due to the satellite transponder. As stated earlier, in **DTH** applications, the receivers are cost-driven, therefore, counter-measures against nonlinearities are applied at the transmitter side. This chapter presents the state-of-the-art of predistortion techniques.

In the literature, the predistortion methods are usually divided into two categories:

Data predistortion which operates at symbol rate (involving \mathcal{F}),

Signal predistortion which operates at higher sampling rate (involving \mathcal{R}).

In this chapter, we adopt another classification depending on whether the predistortion is carried out directly (direct method) or iteratively (iterative method).

The first section is dedicated to direct methods. We start with P -th order inverse method, which is specific to Volterra model, and the global compensation which is compliant with **DVB-S2** standard [41]. We, then, introduce the predistortion technique based on the learning architecture. We end by describing how neural networks apply to predistortion.

The second section focuses on iterative methods, which outperform direct ones. We first detail the small variation algorithm, an iterative data predistortion technique before describing two similar iterative signal predistortion methods, the first based on the fixed-point theorem and the second one which relies on successive approximations.

3.2 Direct predistortion method

3.2.1 P -th order inverse

The p -th order inverse was introduced by Schetzen in [43] and is one of the first predistortion methods for and based on Volterra model. As most systems do not have an inverse or can be inverted only for a limited range of input amplitude, the p -th order inverse is a partial inverse for such systems. All systems, which are **Bounded Input Bounded Output (BIBO)** stable and can be expressed as a *Volterra* series, have a p -th order inverse without input amplitude restriction [37], [43].

The p -th order inverse aims to linearise the system (the first order *Volterra* kernel is a unit impulse response) and cancel *Volterra* kernels from the second to the p -th order. It first requires that the predistortion system is of order p and, secondly, a precise knowledge of the channel.

The cascade, \mathcal{Q} , of the predistortion and of the channel, can be expressed as follows [37], [43]:

$$\mathcal{Q} = \mathcal{I} + \sum_{n=p+1}^{\infty} \mathcal{Q}_n \quad (3.1)$$

The **NISI** correction can either be done at sample or symbol rate.

In [44], two recursive schemes were proposed which allow to reduce the complexity of the synthesis of the predistortion at the cost of higher order operation.

However, the p -th order inverse cannot compensate both the wrapping effect and the clustering effect [45]. Moreover, the p -th order inverse is less efficient than the *Global compensation*, while being also more computationally expensive [45] (see subsection 3.2.2). The *Volterra*-based indirect learning architecture also outperforms the p -th order inverse and does not require the nonlinear model of the channel [46] (see subsection 3.2.3).

In addition, one cannot guarantee significant performance improvement by increasing the order as uncontrolled residual distortions remain.

3.2.2 Global compensation

The *global compensation* algorithm is a data predistortion and exists in two versions: static and dynamic. The static version only corrects the wrapping of the constellation. It is a memoryless method. The dynamic version takes into account previous and future symbols and thus compensates the wrapping and the clustering effect.

The static version was firstly proposed in [47] by Saleh and Salz and the dynamic version was introduced by Karam and Sari in [45].

In order to apply *global compensation* methods, it is unnecessary to model the system since both rely on the *Stochastic gradient method*.

The static version of the algorithm operates as follows. For each transmitted sequence, the centroids (gravity centers) of each symbol of the constellation are computed, then the amplitude and phase error (noted respectively e_ρ and e_θ) from the desired constellation are used to update a **Look-Up Table (LUT)** according to the following relationships:

$$\rho^{k+1}(i) = \rho^k(i) - \alpha_\rho e_\rho(i), \quad 1 \leq i \leq M, \quad (3.2)$$

$$\theta^{k+1}(i) = \theta^k(i) - \alpha_\theta e_\theta(i), \quad 1 \leq i \leq M. \quad (3.3)$$

where $\rho^k(i)$ and $\theta^k(i)$ are respectively the amplitude and the phase of the i -th symbol of the constellation at the k -th iteration.

In the dynamic case, L future and L past symbols are taken into account. Each symbol, then, can take one of $(M+1)^{2L}$ different pre-distorted values. The **LUT** stores $(M+1)^{2L+1}$ values. Update equations have an extra parameter, denoted by l , which corresponds to the index in the **LUT**.

$$\rho^{k+1}(i, l) = \rho^k(i, l) - \alpha_\rho e_\rho(i, l), \quad 1 \leq i \leq M, 1 \leq l \leq (M+1)^{2L}, \quad (3.4)$$

$$\theta^{k+1}(i, l) = \theta^k(i, l) - \alpha_\theta e_\theta(i, l), \quad 1 \leq i \leq M, 1 \leq l \leq (M+1)^{2L}. \quad (3.5)$$

For large constellation, such as 256-**Amplitude and Phase Shift Keying (APSK)**, and long memory effect, this solution might be impractical since the size of the required memory and the time to adapt the predistorted symbols increase exponentially [21].

This problem is partially addressed for **Quadrature Amplitude Modulation (QAM)** by only using the position of a cluster of symbols instead of the accurate future and past symbols. The size of the **LUT** can also be reduced by using the symmetry of the constellation, as in [21], [41] but the solution is still impractical in most cases.

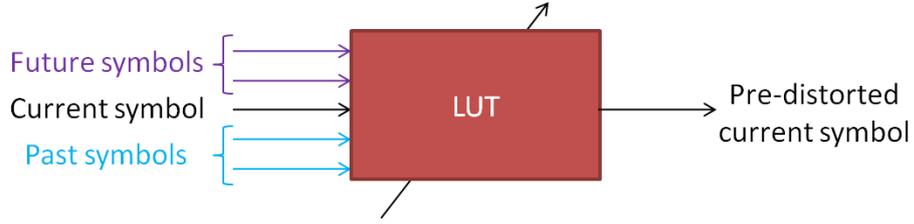


Figure 3.1: Global compensation - Dynamic version

3.2.3 Learning architecture

The P -th order inverse synthesises a Volterra predistorter but it does not guarantee any linearisation improvement due to the remaining high order terms. The direct and indirect learning architecture allow to circumvent this issue by minimizing the quadratic error. The learning architecture aims to adapt a chosen predistortion structure in order to compensate for the nonlinearity. For instance, an Hammerstein pre-distorter [48], [49], a Wiener pre-distorter [50], a Volterra-based pre-distorter [46], [51] and an orthogonal memory polynomial [52] have been trained with an **Indirect Learning Architecture (ILA)**. Likewise with **Direct Learning Architecture (DLA)**, in [53], [54], the predistorter is chosen as a Hammerstein structure to predistort a channel that can be modelled as a Wiener system.

The compensation can either be done at symbol or sample rate.

With an **ILA**, the pre-distorter is a copy of a post-inverse filter that linearises the system. The pre-distorter, then, can be seen as a copy of an equaliser. The adaptive post-inverse filter adapts its weights so that the cascade of the non-linear system and the post-inverse filter minimizes the following cost function:

$$J(l) = \|E(l)\|^2 = \|\zeta\bar{x}_l - z_l\|^2 \quad (3.6)$$

between the input of the channel $\zeta\bar{x}_l$ (of the non-linear system) and the output of the cascade z_l . z is indeed an estimate of the input \bar{x} . The nonlinear predistorter can be adapted by using either an **LMS** algorithm or an **RLS** algorithm [40], [46], [48], [50]–[52].

Figure 3.2 illustrates the indirect learning architecture scheme.

However, since most permutations are irrelevant in the non-linear case and the cost function does not directly minimize the square error $\|\zeta x - y\|^2$, this approach might be less efficient than in the linear case [55].

In [33], in answer to that issue, the authors state that the approach is valid from a theoretical standpoint since as for the p -th order inverse when the system is approximated by a finite-order nonlinear model, the post-inverse and pre-inverse are identical.

The **DLA** adapts the predistortion based on a prior identification of the channel, which leads to a complexity increase to the advantage of better performance with respect to the **ILA** [54]. Figure 3.3 illustrates the direct learning architecture scheme. The *direct learning architecture* has been proposed to enhance the resilience to noise level of the **ILA** [56].

There are two main versions of the **DLA** which are the **Non-linear filtered-X (NFX)** and **Filtered- ϵ (F ϵ)** (also called adjoint filtered-X). Both can be **LMS-**

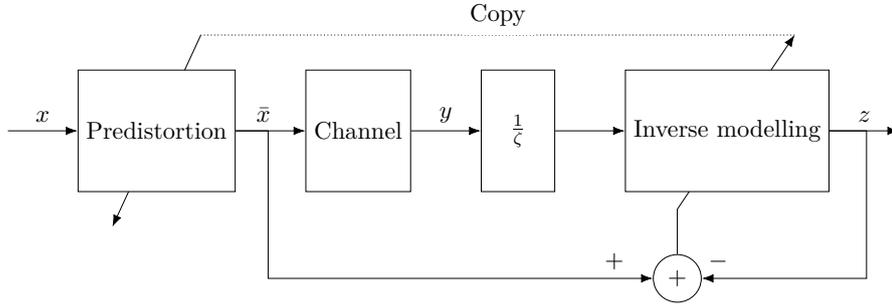


Figure 3.2: Indirect Learning Architecture

based or [RLS](#)-based. The difference between the two comes from the update equation of the pre-distorter.

The performance reachable after convergence is the same but the convergence speed of the $\mathbf{F}\epsilon$ may be slower than the [NFX](#) [54], [57].

In both versions of [DLA](#) a prior identification of the system is required in order to adapt the weights of the predistorter.

The objective function is:

$$J(n) = E\{|e(n)|^2\} \quad (3.7)$$

$$= E\{|\zeta x(n - D) - y(n)|^2\} \quad (3.8)$$

where D is the delay introduced by the system and ζ is the overall system gain.

Nonlinear Filtered-X / DLA

The [Non-linear filtered-X LMS \(NFXLMS\)](#) can only be derived under the assumptions that each function ϕ_i (the i -th component of Φ) is differentiable everywhere except possibly at a few points [54].

The output of the predistorter is computed with the generic Volterra model defined in equation (2.22):

$$\bar{x}_n = \mathbf{W}^T \Phi[\mathbf{X}_n] \quad (3.9)$$

We derive the objective function with respect to the kernels,

$$\frac{\partial J_n}{\partial \mathbf{W}_n^*} = -e(n) \frac{\partial y_n^*}{\partial \mathbf{W}_n^*}. \quad (3.10)$$

The second part of the derivative of the objective function namely $\frac{\partial y(n)}{\partial \mathbf{W}^*(n)}$ is assumed to be negligible in this section.

$$\frac{\partial y^*(n)}{\partial \mathbf{W}^*(n)} = \sum_{l=-L}^L \frac{\partial y^*(n)}{\partial \bar{x}^*(n-l)} \frac{\partial \bar{x}^*(n-l)}{\partial \mathbf{W}^*(n)} \quad (3.11)$$

We assume that the weight evolves slowly over time that is to say $\mathbf{W}(n) \approx \mathbf{W}(n-l)$. This hypothesis is commonly used in the literature.

Given this approximation, we have:

$$\frac{\partial \bar{x}^*(n-l)}{\partial \mathbf{W}^*(n)} \approx \Phi^*[\mathbf{X}(n-l)] \quad (3.12)$$

From eq.(2.23), we obtain:

$$\frac{\partial y(n)}{\partial \bar{x}(n-l)} = Q^T \frac{\partial \Theta[\bar{\mathbf{X}}(n)]}{\partial \bar{x}(n-l)} \quad (3.13)$$

$$= \theta(n, l). \quad (3.14)$$

The update equation is, then, given by:

$$\mathbf{W}(n+1) = \mathbf{W}(n) + \mu e(n) \frac{\partial y^*(n)}{\partial \mathbf{W}^*(n)}. \quad (3.15)$$

By using eq.(3.12), eq.(3.14) and eq.(3.15) becomes:

$$\begin{aligned} \mathbf{W}(n+1) = & \mathbf{W}(n) \\ & + \mu e(n) \sum_{l=-L}^L \theta^*(n, l) \Phi^*[\mathbf{X}(n-l)]. \end{aligned} \quad (3.16)$$

The name of the algorithm comes from the filtered version of the input in the update term of eq. (3.16). The filter $[\theta(n, -L), \dots, \theta(n, 0), \dots, \theta(n, L)]^T$ is called the **Instantaneous Equivalent Linear Filter (IEL)**.

The properties of the **Filtered-X (Fx)** algorithm - especially the robustness against mismatch in the identification, might be lost for the **NFX** since an underlying assumption is the possibility of commutation of blocks, which might be forbidden in a non-linear environment [55, pp.312].

Nonlinear Filtered-error - (F- ϵ)

The **F ϵ** uses a time-delayed time-reverse model of the system to be linearised and has an update equation similar to **Fx** one [57].

As mentioned by Elliott in [57], two versions of the **F ϵ** have been proposed. The difference lies in the objective function. Widrow and Walach in [55] aim to minimise “the mean square value of a filtered error signal, obtained by passing the output error through the inverse of the plant response, whereas the algorithm outlined above [which was the paper contribution] minimises the mean square value of the output error itself” [57].

Generalized DLA - Nonlinear Filtered-X

The generalized **DLA** was proposed in [15] and it is a more complex but more rigorous derivation of the method since it takes into account the second part of the objective function derivative $(\frac{\partial y(n)}{\partial \mathbf{W}^*(n)})$.

The derivative of objective function is, then:

$$\frac{\partial J(n)}{\partial \mathbf{W}(n)^*} = -(e^*(n) \frac{\partial y(n)}{\partial \mathbf{W}^*(n)} + e(n) \frac{\partial y^*(n)}{\partial \mathbf{W}^*(n)}) \quad (3.17)$$

where

$$\frac{\partial y(n)}{\partial \mathbf{W}^*(n)} = \sum_{l=-L}^L \frac{\partial y(n)}{\partial \bar{x}^*(n-l)} \frac{\partial \bar{x}^*(n-l)}{\partial \mathbf{W}^*(n)}, \quad (3.18)$$

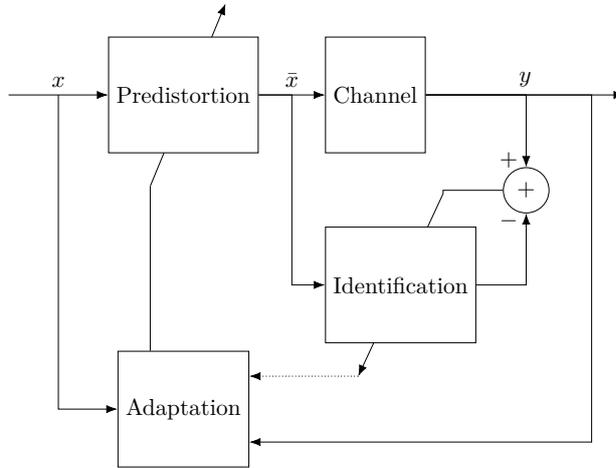


Figure 3.3: Direct Learning Architecture

with

$$\frac{\partial y(n)}{\partial \bar{x}^*(n-l)} = Q^T \frac{\partial \Theta[\bar{\mathbf{X}}(n)]}{\partial \bar{x}^*(n-l)} \quad (3.19)$$

$$= \vartheta(n, l). \quad (3.20)$$

The filter $[\vartheta(n, -L), \dots, \vartheta(n, 0), \dots, \vartheta(n, L)]$ is also an instantaneous equivalent linear filter.

The update equation for the Generalized DLA is thus given by:

$$\begin{aligned} \mathbf{W}(n+1) &= \mathbf{W}(n) \\ &+ \mu \sum_{l=-L}^L \Phi^*[\mathbf{X}(n-l)] \left(e^*(n) \vartheta(n, l) + e(n) \theta^*(n, l) \right). \end{aligned} \quad (3.21)$$

In [15], an adaptive crest factor reduction block which performs a clipping of the signal before the pre-distortion is added. This leads to an improvement of the Signal-to-Interference plus Noise Ratio (SINR) (considering the distortion as a source of noise) of nearly 2dB with respect to the pre-distorter based on DLA alone. The threshold of the clipping is adapted iteratively.

3.2.4 Neural network as a predistortion technique

Neural network has been used for predistortion with or without memory, see for example [58]–[63].

In the context of telecommunication, feedforward neural networks and notably MultiLayer Perceptron (MLP) are used [59]. When the temporal aspect must be taken into account, a short-term memory (tapped delay line) is used to feed feedforward networks, which can be referred to as *Time focused lagged feedforward network* [64].

A layer for a **MLP** is defined by its input/output equation which reads:

$$\mathbf{O}_k = a(\mathbf{W}_k \mathbf{X}_k + \mathbf{B}_k) \quad (3.22)$$

where \mathbf{O}_k is the output vector of the k -th layer, \mathbf{X}_k is the input vector of the k -th layer, \mathbf{W}_k is a weight matrix of the k -th layer, \mathbf{B}_k is the bias vector of the k -th layer and $a(\cdot)$ is the activation function (nonlinear function).

Different types of activation function can be used, as for instance, rectifiers (ReLU) and sigmoids such as hyperbolic tangent.

It has been proved that a **MLP** with a single hidden layer with a finite number of hidden neurons and a sigmoid activation function can approximate any continuous function of any finite number of real variables [65]. Due to their generic aspect, it can be interesting to use them as predistorter [60].

In the case of regenerative satellite, where the predistorter is placed before the amplifier on-board the satellite, the predistorter has, then, to compensate the impairment due to the **HPA** and **OMUX** filter (sometimes, the **OMUX** filter is omitted). A neural network trained with an indirect learning architecture has been proposed in [58], [59] in this context. In the case of the cascade of the amplifier and the **OMUX**, which forms an Hammerstein model, the authors in [59] proposed to use a linear network (which is equivalent to a linear filtering) and a two layers **MLP** is used to compensate the nonlinear part. This is similar to what has been done in [50] for the compensation of an Hammerstein system with a Wiener predistorter (linear filtering followed by a nonlinear element). However, both schemes differ by first the nonlinear element description (polynomial model in [50] and as sum of hyperbolic tangent in [59]) and, secondly, by their learning schemes (backpropagation in [59] and recursive prediction error method in [59]).

The signal predistortion proposed in [61], [62] used a neural network in a similar fashion to predistort a **TWTA** (without **OMUX** and **OMUX**). Likewise, in a communication context, a **MLP** which uses a tapped advance and delay line has been proposed to predistort an amplifier. It outperforms schemes which only rely on the current sample and past samples [60].

Recurrent neural networks, which introduce feedback and, therefore, the output depends on the current state of the neural network [64], were proposed in [63] and were applied in [66] for the compensation of nonlinearities.

Carried out in the context of equalization of a satellite communication channel, simulation reported in [67] shows that complex-valued neural network leads to better performance than real-valued one (where the real and imaginary part are separated).

The drawback with neural network is that the performance of the predistortion is linked to the chosen structure (activation function, number of neurons, number of layers, constraints on the weight matrix, real-valued or complex-based) and in some cases, the convexity cannot be ensured [68]. However, the first applications to predistortion are promising and further investigation is required to overcome this drawback.

3.3 Iterative predistortion methods

In previous section, the signal or data are only processed once. In this section, we focus on iterative methods where the predistortion output is available only when

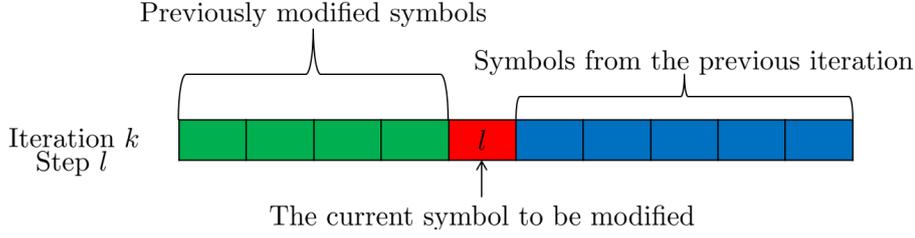


Figure 3.4: Small variation update mechanism

the algorithm has converged. We introduce three kinds of iterative methods. The first one is based on the small-variation algorithm, the second one relies on the fixed point approach and the third one applies successive approximations.

3.3.1 Small variation algorithm

The *small-variation algorithm* proposed by Deleu, Derwin, Kasai, *et al.* in [69], is an iterative algorithm whose convergence is insured by design. This method requires a prior identification of the system.

The algorithm takes as input a block of symbols and aims at minimizing $\|\hat{\mathbf{d}} - \mathbf{d}\|_2$, the Euclidean distance between the received sequence of symbols $\hat{\mathbf{d}}$, and the desired transmitted sequence \mathbf{d} . However, the predistortion is updated symbol-wise.

An iteration refers to a whole block update. It is subdivided into N steps, where N represents the number of symbols in the block. At step l , only the l -th symbol is modified, while all the others remain the same. Figure 3.4 illustrates the update mechanism.

The n -th predistorted symbol $\bar{d}_n^{(k,l)}$ at iteration k and at step $l \neq 1$ is updated according to the following formula:

$$\bar{d}_n^{(k,l)} = \begin{cases} \bar{d}_n^{(k,l-1)} & n \neq l, \\ \arg \min_{\bar{d}_n^{(k,n)}} \|\hat{\mathbf{d}} - \mathbf{d}\|_2 \mid \forall i \neq 1 : d_i = \bar{d}_i^{k,l-1} & n = l. \end{cases} \quad (3.23)$$

First step of iteration k ($l = 1, k \neq 1$) is initialized by the symbol sequence at last step of iteration $k - 1$, and for the first iteration ($k = 1$), the unpredistorted symbols, $\bar{d}_n^{(k-1,N)}$ equals d_n .

$$\bar{d}_n^{(k,1)} = \begin{cases} \bar{d}_n^{(k-1,N)} & n \neq 1, \\ \arg \min_{\bar{d}_1^{(k,1)}} \|\hat{\mathbf{d}} - \mathbf{d}\|_2 \mid \forall i \neq 1 : d_i = \bar{d}_i^{k-1,N} & n = 1. \end{cases} \quad (3.24)$$

To reduce the complexity involved in the computation of equation (3.23), the *small-variation algorithm* gives a suboptimal solution, which is based on the following update:

$$\bar{\mathbf{d}}^{(k,l)} = \bar{\mathbf{d}}^{(k,l-1)} + \mathfrak{d}^{(k,l)} \quad (3.25)$$

with $\mathfrak{d}^{(k,l)}$ is a vector with only one non-zero value, the l -th element, that equals $\bar{d}_l^{(k,l)}$.

The optimal increment $\mathfrak{d}_{opt}^{(k,l)}$ is defined by:

$$\begin{aligned} \mathfrak{d}_{opt}^{(k,l)} &= \arg \min_{\mathfrak{d}^{(k,l)}} [||\mathbf{d} - \mathcal{F}[\bar{\mathbf{d}}^{(k,l-1)} + \mathfrak{d}^{(k,l)}]||_2] \\ &= \arg \min_{\mathfrak{d}^{(k,l)}} [||\underbrace{\mathbf{d} - \mathcal{F}[\bar{\mathbf{d}}^{(k,l-1)}]}_{\mathbf{e}^{(k,l-1)}} - \underbrace{\mathcal{F}[\bar{\mathbf{d}}^{(k,l-1)} + \mathfrak{d}^{(k,l)}] - \mathcal{F}[\bar{\mathbf{d}}^{(k,l-1)}]}_{\mathbf{F}_{NL}^{(k,l)}}}||_2] \\ &= \arg \min_{\mathfrak{d}^{(k,l)}} [||\mathbf{e}^{(k,l-1)} - \mathbf{F}_{NL}^{(k,l)}||_2] \end{aligned} \quad (3.26)$$

where $\mathbf{e}^{(k,l-1)}$ is the error at the previous step and $\mathbf{F}_{NL}^{(k,l)}$ represents the variation of the output. However, the optimal value of $\mathbf{F}_{NL}^{(k,l)}$ is difficult to find. To circumvent this issue, the authors proposed a linear approximation which enables to find a solution but the update is no longer optimal in terms of performance.

The linear approximation reads:

$$\mathbf{F}_{NL}^{(k,l)} \approx \mathbf{F}_{lin}^{(k,l)} = \frac{\partial \mathbf{F}_{NL}^{(k,l)}}{\partial \mathfrak{d}^{(k,l)}} \mathfrak{d}^{(k,l)} + \frac{\partial \mathbf{F}_{NL}^{(k,l)}}{\partial \mathfrak{d}^{*(k,l)}} \mathfrak{d}^{*(k,l)} \quad (3.27)$$

The linear increment is computed as follows:

$$\mathfrak{d}_{lin}^{(k,l)} = \arg \min_{\mathfrak{d}^{(k,l)}} [||\mathbf{e}^{(k,l-1)} - \mathbf{F}_{lin}^{(k,l)}||_2] \quad (3.28)$$

The update equation, then becomes:

$$\bar{\mathbf{d}}^{(k,l)} = \bar{\mathbf{d}}^{(k,l-1)} + \mathfrak{g} \mathfrak{d}_{lin}^{(k,l)} \quad (3.29)$$

with \mathfrak{g} a real damping factor ($\mathfrak{g} \in [0, 1]$) that ensures the convergence of the method. The damping factor \mathfrak{g} sets the trade-off between the quality and the speed of the convergence.

In [69], the method was shown to outperform dynamic global compensation with a memory depth of 3 and memory polynomial-based indirect learning architecture.

However, this small-variation algorithm is fairly complex as it requires N steps per iteration and each step involves the computation of the linear approximation $\mathbf{F}_{lin}^{(k,l)}$ and the minimization of $||\mathbf{e}^{(k,l-1)} - \mathbf{F}_{lin}^{(k,l)}||_2$.

To reduce the complexity, a raw model of the channel is used during the steps within an iteration in order to compute the update. At the end of each iteration, a finer (more accurate) model is applied to compute the stopping criterion. If the error $\mathbf{e}^{(k,N)}$ decreased then the iterative process goes on. Otherwise, the pre-distorted sequence of the previous iteration is used as predistortion output and the process is over.

Moreover, the performance is conditioned by the choice of the damping factor \mathfrak{g} .

3.3.2 Iterative predistortion based on the fixed-point theorem

Iterative predistortion technique based on the contraction mapping theorem has been proposed to linearize a non-linear channel either with memory [18], [70],

[71] or without memory [72]. The predistortion problem can be formulated as follows:

$$\mathcal{R}[\bar{\mathbf{X}}(l)] = \zeta \mathbf{X}(l), \quad (3.30)$$

where ζ defines the desired ideal gain at the output of the non-linear channel.

The value of ζ should be taken such that $\zeta \mathbf{X}(l)$ belongs to the possible set of output of the non-linear channel. For instance, the output power cannot be higher than the maximum power that can be delivered by the amplifier.

The problem expressed by equation (3.30) can be reformulated as a fixed-point problem, either by inverting the linear part of \mathcal{R} as in [70], [71] or by adding on both sides $(I - \mathcal{R})[\bar{\mathbf{X}}(l)]$ as in [18]. The latter solution avoids a computationally expensive inversion of the linear part of the channel.

The mathematical formulation of the predistortion problem as a fixed-point one is then:

$$(I - \mathcal{R})[\bar{\mathbf{X}}(l)] + \zeta \mathbf{X}(l) = \bar{\mathbf{X}}(l). \quad (3.31)$$

If the conditions for the contraction mapping theorem are met for the operator $\mathcal{T}[\cdot] = (I - \mathcal{R})[\cdot] + \zeta \mathbf{X}(l)$, the solution is unique and it can be iteratively reached by using:

$$\bar{\mathbf{X}}^{(k+1)}(l) = \mathcal{T}[\bar{\mathbf{X}}^{(k)}(l)], \quad (3.32)$$

where k is the stage number of the iterative process. The conditions for the operator $\mathcal{T}[\cdot]$ to be a contraction mapping, when the non-linear system can be described as a Volterra model, are given in [18], [70]. However, the verification of the conditions might be cumbersome as they depend on the chosen norm. Often, the conditions cannot be computed and upper bounds are used [70]. On the other hand, if the conditions are unfulfilled, performing few iterations might improve the linearity of the equivalent channel [18]. Moreover, an appropriate constant γ can be introduced to make $\mathcal{T}[\cdot]$ a contraction mapping [72]. The recursion equation (3.32) then becomes:

$$\bar{\mathbf{X}}^{(k+1)}(l) = \bar{\mathbf{X}}^{(k)}(l) + \gamma \left(\zeta \mathbf{X}(l) - \mathcal{R}[\bar{\mathbf{X}}^{(k)}(l)] \right). \quad (3.33)$$

γ has to be tuned empirically for each channel operator \mathcal{R} . The parameter γ can be used to match the output range of the model with its input range. Moreover, this parameter affects the speed and the quality of the convergence. In some cases, it should be made decreasing with the iterations, which requires further empirical tuning [73].

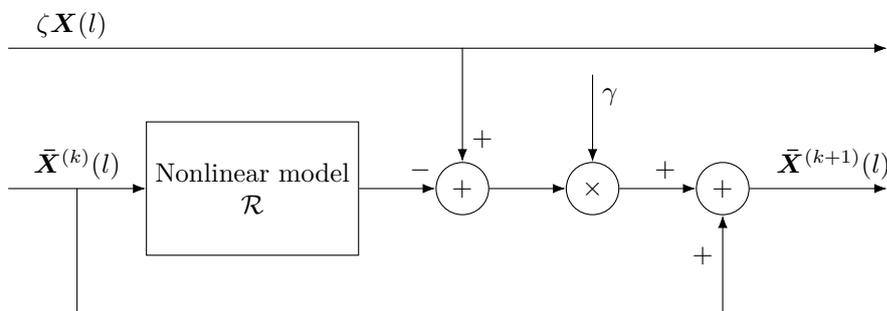
Figure 3.5 illustrates the k -th stage of the iterative process.

The complexity of this structure is mainly driven by the nonlinear model used for the computation of update.

3.3.3 Successive approximation

A similar structure to the iterative structure of the one derived through the lens of the fixed-point theorem is presented in [30].

The two main differences are the introduction of complex-valued gain correction and a step-size that helps to ensure convergence. Given the model output, the gain correction enables to subtract its projection on the reference signal and thus retain the orthogonal part only. The step-size, therefore, is blending the

Figure 3.5: k -th stage of the iterative signal predistortion scheme.

orthogonal part of the model's output with the predistorted signal computed at the previous iteration.

Three versions are proposed. The first one, given in Figure 3.6, focuses on the suppression of the spectral regrowth. The second one, illustrated in Figure 3.7 aims at reducing the in-band distortion. Finally, the third one tries to meet a trade-off between in-band and out-of-band distortion. It is summarized in Figure 3.8. It can be seen as a combination of the two first schemes. Two combining factors (α and β) are used to balance the amount of suppression of the spectral regrowth with respect to the removal of the in-band distortion.

In the first two schemes, the update equation reads:

$$\bar{\mathbf{X}}_m^{(k+1)}(l) = \bar{\mathbf{X}}_m^{(k)}(l) + \mu_m^{(k)} \mathbf{E}_m(l), \quad (3.34)$$

with $m = 1$ or 2 depending on the scheme (1 - Out-band distortion suppression, 2 - In-band distortion) and k is the stage number of the iterative process.

The update error is defined as:

$$\mathbf{E}_1(l) = \mathbf{X}(l) - \gamma_1^{-1} \mathcal{R}\{\bar{\mathbf{X}}^{(k)}(l)\}, \quad (3.35)$$

for the first scheme and

$$\mathbf{E}_2(l) = \mathbf{G}_{Rx}^{up} \left(\mathbf{X}(l) - \gamma_2^{-1} \mathcal{R}\{\bar{\mathbf{X}}^{(k)}(l)\} \right), \quad (3.36)$$

for the second one, with \mathbf{G}_{Rx}^{up} is a matrix generated based on the receiver's filter without downsampling, so \mathbf{G}_{Rx}^{up} is a square matrix ($M \times M$).

γ_1 and γ_2 ¹ are coefficients that aim to correct the wrapping at the output of the non-linear model. They are, respectively, defined by:

$$\gamma_1^{(k)} = \frac{\mathcal{R}\{\bar{\mathbf{X}}^{(k)}(l)\}^* \mathbf{X}(l)}{\|\mathbf{X}(l)\|^2}, \quad (3.37)$$

and

$$\gamma_2^{(k)} = \frac{(\mathbf{G}_{Rx} \mathcal{R}\{\bar{\mathbf{X}}^{(k)}(l)\})^* (\mathbf{G}_{Rx} \mathbf{X}(l))}{\|\mathbf{G}_{Rx} \mathbf{X}(l)\|^2}. \quad (3.38)$$

Those gains help to only retain the orthogonal part of the reference signal. However, they introduce a scaling inversely proportional to the correlation between the reference signal and the output of the model.

¹The exact formulation of γ_2 is not given in [30] but has a textual definition similar to γ_1 .

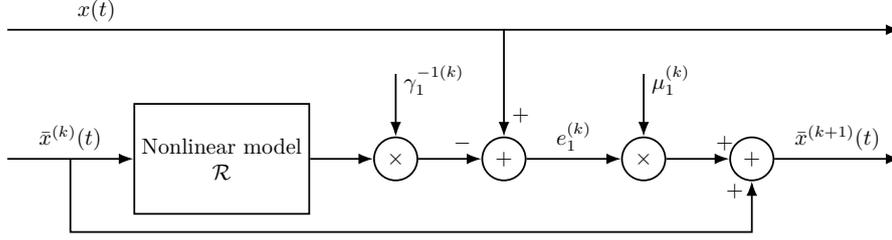


Figure 3.6: Successive approximation - First scheme - Suppression of the spectral regrowth [30]

The step-size sequence $\{\mu_m^{(k)}\}_{m=0}^{\infty}$ (the step-sizes are denoted by μ_1 and μ_2 in figures 3.6 to 3.8) must satisfy the following conditions:

$$\mu_m^{(k)} > 0 \quad (3.39)$$

$$\lim_{k \rightarrow \infty} \mu_m^{(k)} = 0 \quad (3.40)$$

$$\sum_{k=0}^{\infty} \mu_m^{(k)} = \infty \quad (3.41)$$

$$\sum_{k=0}^{\infty} (\mu_m^{(k)})^2 < \infty \quad (3.42)$$

The convergence should be insured, as long as the chosen step-size fulfils conditions given in equations (3.39) to (3.42). It isn't necessary to take into account the characteristics of the non-linear system or of the input signal.

A general expression that satisfies equations (3.39) to (3.42) is given in [30]:

$$\mu^{(k)} = \mu^{(0)} \frac{\left(\frac{b}{k+1} + a\right)}{\left(a + \frac{b}{k+1} + (k+1)^c - 1\right)}, \quad (3.43)$$

where the quadruplet $(\mu^{(0)}, a, b, c)$ has to be optimized.

The update equation for the third scheme is given by:

$$\bar{\mathbf{X}}^{(k+1)}(l) = \bar{\mathbf{X}}^{(k)}(l) + (\alpha\mu_1^{(k)}\mathbf{E}_1(l) + \beta\mu_2^{(k)}\mathbf{E}_2(l)) \quad (3.44)$$

The nonlinear model is identified and is updated during off-line training phase. The training phase should be done either when a change in the environment has occurred or periodically to track variation in the non-linear system characteristics. The complexity is, then, mainly driven by the non-linear model.

3.4 Conclusion

The purpose of this chapter was to deliver a state-of-the-art of predistortion techniques which apply either to the data or to the signal.

We first introduced direct methods that compute the predistorted sequence in one pass. Among them let us mention the Volterra predistorter trained with direct and indirect learning architecture, and neural-network based techniques which were not studied in this PhD but could be one of the perspectives of

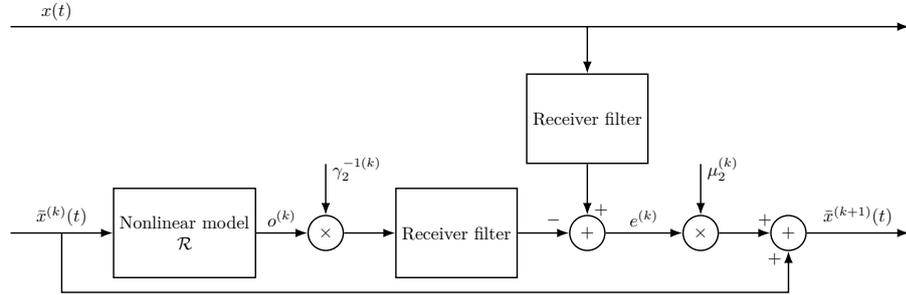


Figure 3.7: Successive approximation - Second scheme - Suppression of the in-band distortion [30]

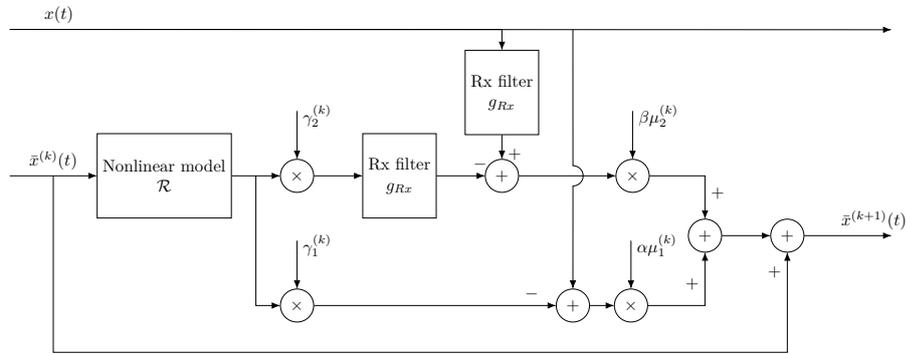


Figure 3.8: Successive approximation - Third scheme - Combination of the two scheme (suppression of spectral regrowth and in-band distortion)[30]

this work. We, then, focused on iterative methods which are known to outperform direct methods but require empirical tuning of parameters so as to ensure both convergence and performance. Among selected techniques, let us mention, the fixed-point theorem based predistortion which has the advantage of a well-defined structure and good performance but requires that the contraction mapping theorem condition is satisfied. In that case, the structure of the predistorter is directly induced by the identification. This removes the need for expertise regarding the choice of a predistorter model. With a structure derived from the fixed-point theorem, improving the identification leads to an improvement of the linearisation performance. In the following chapter, we start from this scheme and direct learning architecture to derive iterative predistortion schemes that achieve good performance in terms of in-band and out-of-band distortion mitigation as well as total degradation and that do not require empirical tuning of parameters and should converge for a large class of satellite transponders.

Chapter 4

Iterative predistortion

“Is it wonderful ?”

Kalopsia - ...Like Clockwork
Queens of the stone age
Matador Records. 2013

Contents

4.1 Introduction	62
4.2 Data-optimized signal predistortion	62
4.2.1 Data predistortion based on the contraction mapping theorem	62
4.2.2 Data-optimized signal predistortion scheme	64
4.2.3 Comparison of the data-optimized signal predistortion with the successive approximation scheme assuming perfect identification	64
4.2.4 Comparison with signal predistortion based on the contraction mapping	66
4.3 Iterative adaptive signal predistortion	69
4.3.1 Iterative adaptive signal-optimized predistortion	70
4.3.2 Iterative adaptive data-optimized signal predistortion	70
4.3.3 Adaptation process	71
4.3.4 Comparison with signal-optimized signal predistortion scheme	74
4.3.5 Robustness towards feedback mismatch	76
4.4 Perspectives	76
4.4.1 Backpropagation	76
4.4.2 Joint Clipping and predistortion	79
4.4.3 Variable step-size	79
4.5 Conclusion	80

4.1 Introduction

At the end of the overview done in the previous chapter, we concluded that iterative predistortion methods based on the fixed-point theorem are well suited to deal with nonlinearities. Firstly, iterative methods outperform direct ones. Secondly, they don't require expertise to select a predefined predistortion structure and only need the channel model. Thirdly, when the conditions of the contraction mapping theorem are satisfied, the iterative algorithm converges towards the optimum fixed-point without requiring empirical parameter tuning.

In the first part of the chapter, we focus on data predistortion based on the contraction mapping theorem, by taking into account low-pass transmit and receive filters, since data predistortion has been proposed to be used with [DVB-S2](#) modem in [\[41\]](#). Moreover, data predistortion does not increase the spectral support required on the uplink. However, they are unable to compensate the spectral regrowth. We, then, extend the proposed scheme to define a data-optimized signal predistortion scheme which performs the same as the original data predistortion one. One of the advantages of using a signal predistortion is to avoid the need to identify the transmit and receiver filter, which reduces the length of the impulse response.

In practice, there is a trade-off to be met between residual distortion and [OBO](#) which may question the validity of the contraction mapping theorem's premises. A solution may be the addition of a damping factor that reduces the value of the update and thus leads to the same drawback as most iterative methods. To circumvent this issue, inspired by the [DLA](#), we introduce a double adaptive scheme optimized so as to ensure either the optimum convergence or the best performance for a target number of iterations.

4.2 Data-optimized signal predistortion

In this section, we first derive the data predistortion scheme based on the contraction mapping theorem. Then, we derive a signal version of the proposed method, which can be used with a signal model of the satellite transponder.

The results presented in that section have been published in:

- N. Alibert, K. Amis, C. Langlais, D. Castelain. "Signal predistortion scheme based on the contraction mapping theorem", *GRETSI 2017 : 26ème colloque du Groupement de Recherche en Traitement du Signal et des Images*, 2017, Juan-les-Pins, France.
- N. Alibert, K. Amis, C. Langlais, D. Castelain. "Comparison of signal predistortion schemes based on the contraction mapping for satellite communications with channel identification", *ICT 2018: 25th International Conference on Telecommunications*, Saint Malo, France.

4.2.1 Data predistortion based on the contraction mapping theorem

Predistortion scheme based on the fixed-point theorem and available in the literature belongs to the signal predistortion class (see section [3.3.2](#)). In this

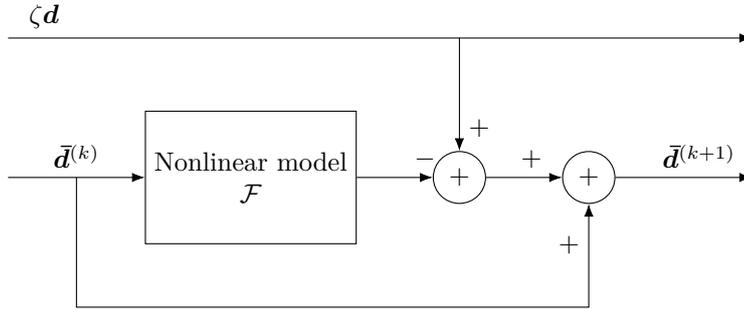


Figure 4.1: Data predistortion structure

section, we propose a data predistortion scheme optimized according to the NMSE criterion.

Firstly, we write the problem of predistortion for a data predistortion. We look for a predistorted sequence $\bar{\mathbf{d}}$ which satisfies:

$$\mathcal{F}[\bar{\mathbf{d}}] = \zeta \mathbf{G}_{Rx} \mathbf{H}_{Tx} \mathbf{d}. \quad (4.1)$$

with \mathcal{F} , \mathbf{G}_{Rx} and \mathbf{H}_{Tx} defined in Section 2.2.

For clarity of the derivation, in the following, we assume that h_{Tx} and g_{Rx} (and, therefore, \mathbf{H}_{Tx} and \mathbf{G}_{Rx}) satisfy the Nyquist criterion. Thus, $\mathbf{G}_{Rx} \mathbf{H}_{Tx} = \mathbf{I}$, with \mathbf{I} being the identity matrix.

For instance, with **Faster-than-Nyquist (FTN)** pulse shaping filters, the convolution of the transmit and receive filters introduces inter-symbol interferences [74], [75]. In similar cases, it would only imply a matrix multiplication of the reference with $\mathbf{G}_{Rx} \mathbf{H}_{Tx}$.

Then, following similar derivation to the one done in [18], [70] (see section 3.3.2), equation (4.1) is re-written as a fixed-point problem formulation:

$$(\mathbf{I} - \mathcal{F})[\bar{\mathbf{d}}] + \zeta \mathbf{d} = \bar{\mathbf{d}}, \quad (4.2)$$

where \mathcal{I} is the identity operator.

If the operator $\mathcal{T}[\cdot] = (\mathbf{I} - \mathcal{F})[\cdot] + \zeta \mathbf{d}$ is a contraction mapping, then, according to the contraction mapping theorem, the solution is unique and can be iteratively reached by using [70], [72]:

$$\bar{\mathbf{d}}^{(k+1)} = \mathcal{T}[\bar{\mathbf{d}}^{(k)}]. \quad (4.3)$$

The conditions for the operator \mathcal{T} to be a contraction mapping are the same as the ones defined in section 3.3.2. Therefore, the value of ζ should be taken such that $\zeta \mathbf{d}$ belongs to the possible set of output of the non-linear channel. However, if the hypotheses are not satisfied, a few iterations might still improve the performance [18]. Otherwise, as in [72], a constant step-size can be introduced to ensure that \mathcal{T} is a contraction mapping.

In the end, the recursion can be written as follows:

$$\bar{\mathbf{d}}^{(k+1)} = \bar{\mathbf{d}}^{(k)} + (\zeta \mathbf{d} - \mathcal{F}[\bar{\mathbf{d}}^{(k)}]). \quad (4.4)$$

Figure 4.1 shows the stage for symbol predistortion. We can see that the structure is analogue to the one proposed in the state-of-the-art (see Figure 3.5).

4.2.2 Data-optimized signal predistortion scheme

The data predistortion scheme described in the previous section requires to identify \mathcal{F} which may involve more kernels in the Volterra Model than \mathcal{R} since \mathcal{F} includes transmit and receive low-pass filters in order to compare the proposed method with the state-of-the-art method with the same channel. In this section, we aim to derive the signal predistortion scheme equivalent to the data predistortion scheme in that both optimize the NMSE criterion applied to the data. We first establish the connection between both models.

According to the model introduced in Chapter 2 and illustrated in figure 2.2, the relation between the sampled modulated signal and the symbol sequence \mathbf{d} is given by:

$$\mathbf{X} = \mathbf{H}_{Tx} \mathbf{d}, \quad (4.5)$$

and the data model output can be written as a function of the signal model output:

$$\mathcal{F}[\mathbf{d}] = \mathbf{G}_{Rx} \mathcal{R}[\mathbf{X}]. \quad (4.6)$$

Using (4.4), the equivalence between the signal and the data predistortion schemes reads:

$$\bar{\mathbf{X}}^{(k+1)}(l) = \mathbf{H}_{Tx}(\bar{\mathbf{d}}^{(k)}(l) + \zeta \mathbf{d}(l) - \mathcal{F}[\bar{\mathbf{d}}^{(k)}(l)]) \quad (4.7)$$

$$= \bar{\mathbf{X}}^{(k)}(l) + \zeta \mathbf{X}(l) - \mathbf{H}_{Tx} \mathbf{G}_{Rx} \mathcal{R}[\bar{\mathbf{X}}^{(k)}(l)] \quad (4.8)$$

$$= \bar{\mathbf{X}}^{(k)}(l) + \mathbf{E}^{(k)}(l) \quad (4.9)$$

yielding the error signal:

$$\mathbf{E}^{(k)}(l) = \mathbf{H}_{Tx}(\zeta \mathbf{d}(l) - \mathcal{F}[\bar{\mathbf{d}}^{(k)}(l)]) \quad (4.10)$$

$$= \zeta \mathbf{X}(l) - \mathbf{H}_{Tx} \mathbf{G}_{Rx} \mathcal{R}[\bar{\mathbf{X}}^{(k)}(l)] \quad (4.11)$$

The product $\mathbf{H}_{Tx} \mathbf{G}_{Rx}$ is not the identity matrix and contains multiple columns of zeros, transcribing the sampling operation.

As we can see the update is only composed of the error. The update might be sub-optimal, in comparison of the method proposed by [69]. However, all the symbols are updated at each step, which reduces the overall complexity of the method.

The successive approximation scheme proposed in [30] also minimizes the NMSE. The main difference between both schemes comes from the error computation. In the successive approximation scheme, only the orthogonal part of the upsampled and filtered signal is used to update the predistortion signal, whereas the update is applied on the downsampled signal in our proposed scheme. Therefore, the spectral support of the uplink might be widened in the case of successive approximation.

4.2.3 Comparison of the data-optimized signal predistortion with the successive approximation scheme assuming perfect identification

In this section, we compare the data-optimized signal predistortion aiming at minimizing the NMSE with the second scheme of the successive approximation proposed in [30], which aims to minimize the in-band distortion (see section 3.3.3) through Monte Carlo simulations. This comparison is carried out

in an ideal setup where the nonlinear model of the channel is assumed to be known.

Let us describe the simulations parameters.

Simulation setup

The transmitter and receiver filter are square-root raised cosine (SRRC) filters with a 5% roll-off. The IMUX, OMUX and the TWTA characteristics are the one defined in the DVB-S2 [11]. The symbol rate is set to 38 MBd. The interferers are delayed and time-shifted versions of the output of the OMUX filter located at 40 MHz on both sides of the carrier of interest (see equation (2.2)).

Simulations results

In the simulations, the successive approximation (second scheme) [30] is applied with a genie-aided amplitude and phase correction at the receiver and we consider two step-size sequences defined by (3.43) with $\mu^{(0)} = 1$. $(a, b, c) = (1, 0, 0)$ (yielding $\mu^{(k)} = 1$) for the first one and $(a, b, c) = (10, 0, 0.6)$ for the second one. In the literature, centroid-based receivers are often used, however, as they mitigate for the remaining wrapping uncompensated by the predistorter.

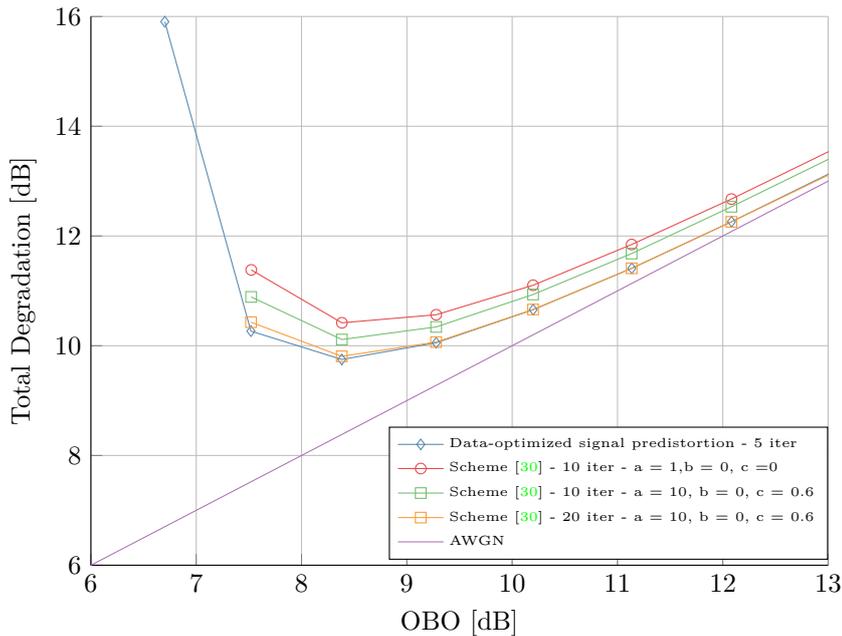


Figure 4.2: Total degradation for a target BER of 10^{-4} .

The TD and the NMSE are compared in Fig. 4.2 and Fig. 4.3 respectively.

Concerning the TD, the performance of the successive approximation method depends on the chosen step-size parameters (eq. (3.43)) and requires more iterations (20 ite.) than ours (5 ite.) to obtain equivalent results.

As for the NMSE, the proposed data-optimized signal predistortion scheme requires less than 5 iterations to converge when the reference scheme needs more

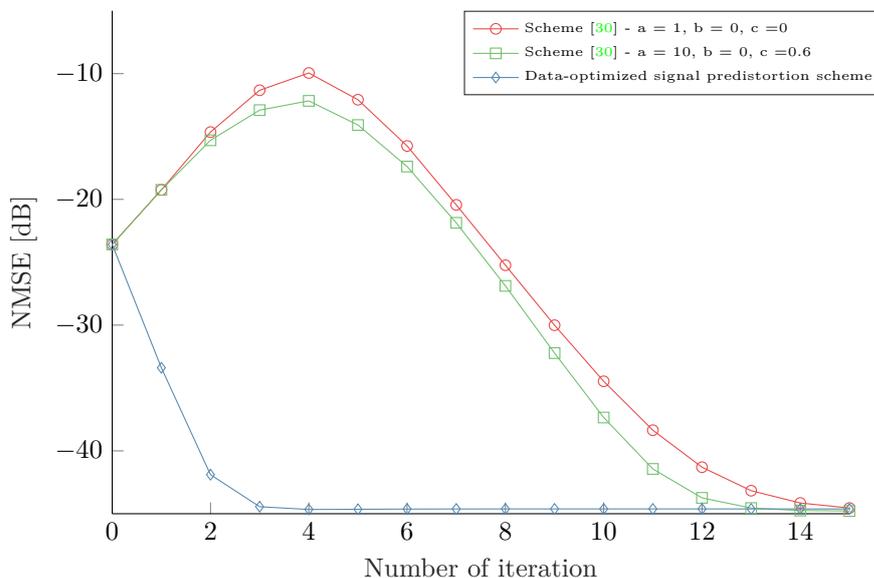


Figure 4.3: NMSE at the output of the matched filter for an IBO of 14 dB versus the number of stages

than 10 iterations to achieve the same NMSE.

Conclusion

The proposed data-optimized signal predistortion scheme outperforms the successive approximation method introduced in [30] in terms of TD and convergence rate. Moreover, contrary to the reference scheme, it does not require defining parameters nor empirical tuning of those parameters when the conditions of the contraction mapping theorem are met. However, those results only consider the reduction of in-band distortion and therefore, only one scheme (the second scheme) of the successive approximation proposed in [30] was considered.

4.2.4 Comparison with signal predistortion based on the contraction mapping

In this section, we compare the data-optimized signal predistortion scheme to the signal predistortion based on the fixed-point theorem, that we will refer as signal-optimized signal predistortion, which was detailed in section 3.3.2 in the case where the channel is either identified or known.

The simulation setup is the same as the one used in section 4.2.3.

Identification procedure

To identify the non-linear system \mathcal{R} , we use a memory polynomial model (MP) [33], which is a reduced-complexity Volterra model [76]. It allows practical on-the-fly implementation [30]. We apply a LMS algorithm, as in [77], with the step parameter set to 10^{-5} and a training sequence that spans 500000 symbols.

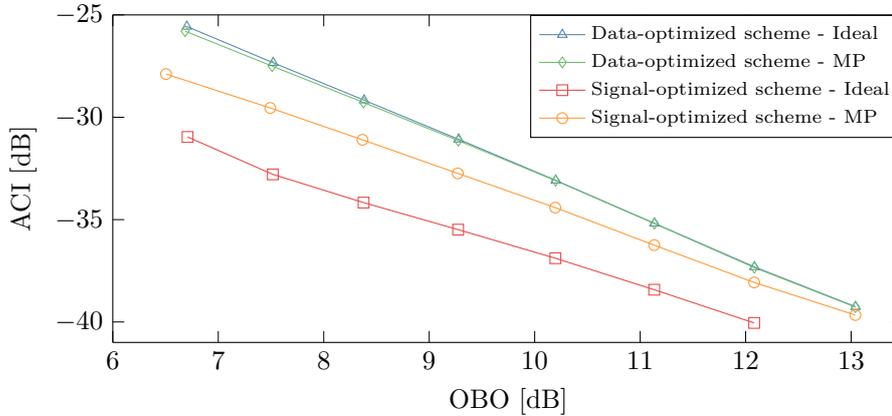


Figure 4.4: ACI at the output of the matched filter versus the OBO.

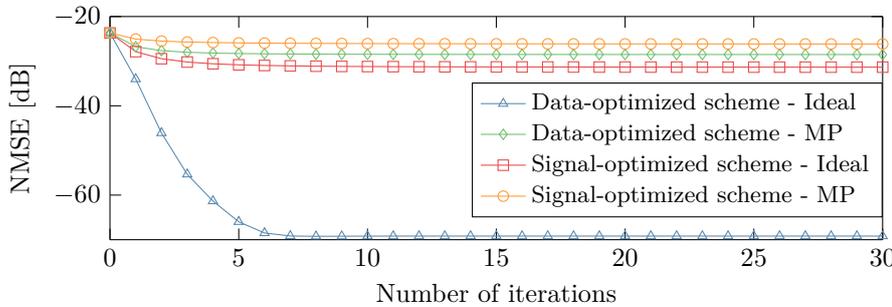


Figure 4.5: NMSE at the output of the matched filter for an IBO of 14 dB versus the number of stages.

As in [76], only odd power terms up to the fifth order ($K = 5$) are considered. The memory depth is set to 2 ($L = 2$).

The same identified channel is assumed for both schemes.

Tuning of the empirical parameter

The empirical parameter is tuned with the bisection method aiming to ensure the convergence. As the parameter γ affects the convergence and the quality of the convergence, we look for the biggest value which ensures the convergence. The step-size parameter is tuned numerically over 1000 iterations with a tolerance of 10^{-4} . We check that the NMSE is decreasing or constant over the 1000 iterations. The NMSE is computed over the real channel and not the identified one.

Results

The ACI performance is given in Fig. 4.4 in both perfect and identified non-linear system model. The signal-optimized signal predistortion scheme outperforms the data-optimized scheme up to 5.4 dB in both cases. However, whereas the data-optimized scheme exhibits similar performance in both cases,

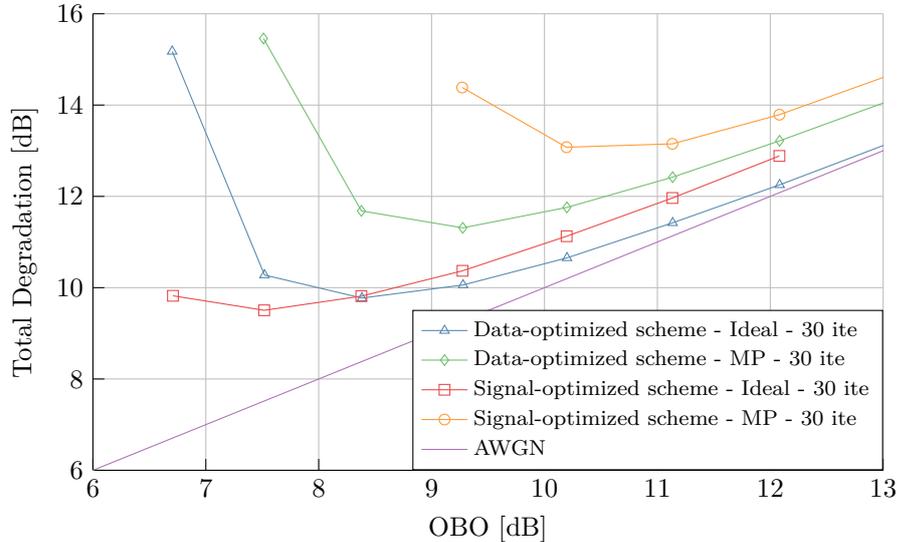


Figure 4.6: Total degradation with ACI for a target BER of 10^{-4} for the data-optimized and the signal-optimized signal predistortion scheme with perfect knowledge of the channel (Ideal) or with identification (MP).

the signal-optimized scheme is more sensitive to identification mismatch (degradation of up to 3.3 dB).

As for the NMSE, given in Fig. 4.5 the data-optimized scheme significantly outperforms the signal-optimized one in both cases (38 dB for the ideal case and 2.5 dB in the identified case). Both methods have converged after 10 iterations.

The total degradation TD *with ACI* is plotted in Fig. 4.6 as a function of OBO. We observe that, at their optimum point, the signal-optimized scheme performs slightly better (0.3 dB) than the data-optimized one when the non-linear channel is perfectly known. The trend is reversed in case of imperfect identification with MP model. The data-optimized scheme largely outperforms the signal-optimized scheme by 1.8 dB.

Conclusion

We conclude that the data-optimized signal predistortion scheme is more effective in practice. The complexities of both methods are nearly equivalent since they are mainly driven by the non-linear channel model and the number of stages of the iterative structure. The data-optimized scheme outperforms the signal-optimized one in terms of TD, NMSE and robustness towards identification mismatch.

When the conditions for a contraction mapping are not met, one can use an empirical gain to make the operator \mathcal{R} a contraction. Moreover, the identification mismatch also impacts the performance and the improvement of the linearization performance cannot be guaranteed. In the following, to circumvent this issue, we propose a self-adaptive scheme whose parameters are automatically adapted.

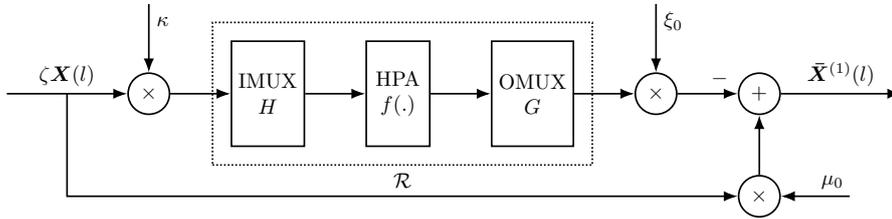


Figure 4.7: First stage of the proposed iterative adaptive signal predistortion scheme.

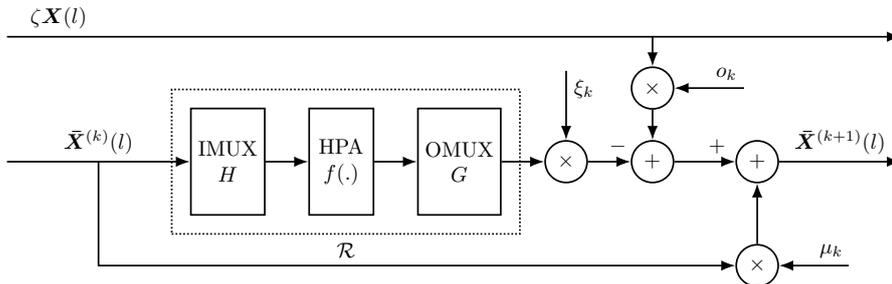


Figure 4.8: k -th stage ($k \geq 1$) of the proposed iterative adaptive signal predistortion scheme.

4.3 Iterative adaptive signal predistortion

We start from the previous formulation of the fixed-point problem given by equation (3.30). In order to circumvent the issue of empirical tuning of parameter, we introduce gains that can be adapted with a feedback from the channel. Those extra gains are used to ensure the best performance after a given number of iterations and convergence when the conditions of the contraction mapping theorem are unfulfilled. Moreover, the working point that minimizes the TD has often some remaining distortion in favour of a smaller OBO. Therefore, this point is out of the scope of convergence of the fixed-point theorem. Like the direct learning architecture [15], [54], the resulting predistortion scheme is twice-adaptive. It requires the channel identification as preliminary step, then the stages are adapted. In this section, we consider that the channel model \mathcal{R} is known.

In this section, we first present the steady-state signal and data-optimized adaptive signal predistortion schemes, then we introduce the adaptation process.

The signal-optimized signal predistortion method has been presented in:

- N. Alibert, K. Amis, C. Langlais, D. Castelain. "Iterative adaptive signal predistortion for satellite communications", ISTC 2018: International Symposium on Turbo Codes & Iterative Information Processing 2018, Hong Kong, Hong Kong.

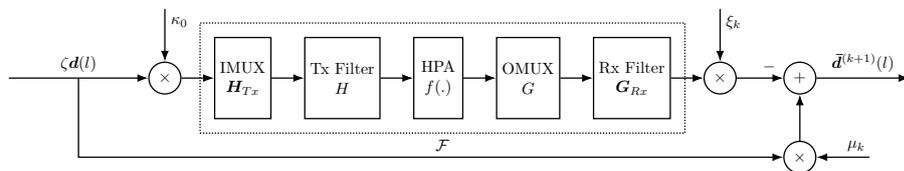


Figure 4.9: First stage of the data-optimized iterative adaptive signal predistortion scheme.

4.3.1 Iterative adaptive signal-optimized predistortion

Based on (3.33), we propose a novel formulation of the iterative signal predistortion scheme such that the predistorted signal at stage k is computed as follows:

$$\bar{\mathbf{X}}^{(k+1)}(l) = \mu_k \bar{\mathbf{X}}^{(k)}(l) + \mathbf{E}^{(k)}(l), \quad (4.12)$$

with $\mathbf{E}^{(k)}(l)$ the update error defined by

$$\mathbf{E}^{(k)}(l) = o_k \zeta \mathbf{X}(l) - \xi_k \mathcal{R}\{\bar{\mathbf{X}}^{(k)}(l)\}, \quad k \geq 1. \quad (4.13)$$

Thus, at the k -th stage ($k \geq 1$), the predistorted signal $\bar{\mathbf{X}}^{(k+1)}(l)$ is computed as the sum of two terms. The first one is the predistorted signal $\bar{\mathbf{X}}^{(k)}(l)$ computed at previous stage. Its contribution to the definition of $\bar{\mathbf{X}}^{(k+1)}(l)$ is weighted by the feedforward gain denoted by μ_k . The second one, the update signal $\mathbf{E}^k(l)$ is the error between the output of the channel model, with the predistorted signal equal to $\bar{\mathbf{X}}^{(k)}(l)$ as input weighted by ξ_k , and the reference signal $\mathbf{X}(l)$, weighted by o_k . Instead of having one parameter weighting the error, we have introduced an extra degree of freedom by using two separate gains. The post model gain, denoted by ξ_k , enables to adjust the level at the output of the channel model. The reference gain, denoted by o_k , aims at controlling the contribution of the reference signal. The combination of the three gains (o_k, μ_k, ξ_k) allows the process to select the amplifier's operating point and blend the three components to generate the output of the stage $\bar{\mathbf{X}}^{(k+1)}(l)$.

At $k = 0$, the process is initialized with $\bar{\mathbf{X}}^{(0)}(l) = \zeta \mathbf{X}(l)$. Equation (4.12) can be simplified since the initialization signal and the reference signal are equal in this case. The first predistorted signal is then expressed as follows:

$$\bar{\mathbf{X}}^{(1)}(l) = \mu_0 \zeta \mathbf{X}(l) - \xi_0 \mathcal{R}\{\kappa \zeta \mathbf{X}(l)\}, \quad k = 0. \quad (4.14)$$

An additional pre-model gain κ is introduced to adjust the operating point of the model depending on the initial signal amplitude (which can be either too low or high). The feedforward gain μ_0 controls the part of the initialization for the next stage and the computation of the error.

Figure 4.7 and 4.8 illustrate first ($k = 0$) and intermediate ($k \geq 1$) stages respectively.

4.3.2 Iterative adaptive data-optimized signal predistortion

In a similar fashion, we derive the data-optimized signal predistortion based on equation (4.11).

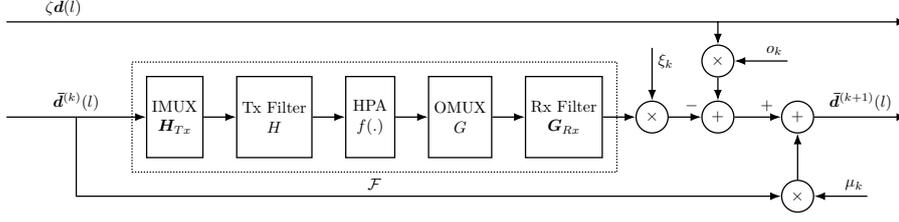


Figure 4.10: k -th stage ($k \geq 1$) of the data-optimized iterative adaptive signal predistortion scheme.

The symbol-optimized predistorted signal at stage k is computed as follows:

$$\bar{\mathbf{d}}^{(k+1)}(l) = \mu_k \bar{\mathbf{d}}^{(k)}(l) + \mathbf{E}^{(k)}(l), \quad (4.15)$$

with $\mathbf{E}^{(k)}(l)$ the update error defined by

$$\mathbf{E}^{(k)}(l) = o_k \zeta \mathbf{d}(l) - \xi_k \mathcal{F}\{\bar{\mathbf{d}}^{(k)}(l)\}, \quad k \geq 1. \quad (4.16)$$

As we initialize the iterative process with $\bar{\mathbf{d}}^{(0)}(l) = \zeta \mathbf{d}(l)$, the update equation of the first iteration can be simplified as follows:

$$\bar{\mathbf{X}}^{(1)}(l) = \mu_0 \zeta \mathbf{d}(l) - \xi_0 \mathcal{F}\{\kappa \zeta \mathbf{d}(l)\}, \quad k = 0. \quad (4.17)$$

The additional gains denoted by κ and μ_0 play the same role as κ and μ_0 in (4.17). Figure 4.9 and 4.10 illustrate first ($k = 0$) and intermediate ($k \geq 1$) stages respectively.

4.3.3 Adaptation process

The adaptation process is only derived for the signal-optimized process, the extension to the data-optimized scheme is straightforward.

The adaptation process consists in defining the gains $(\kappa, o_k, \mu_k, \xi_k)$ through an iterative process that computes $(\kappa_n, o_{k,n}, \mu_{k,n}, \xi_{k,n})$. Subscript k and n refer to the predistortion stage and to the adaptation process iteration, respectively.

The adaptation is performed before the transmission and during the transmission the gains are fixed. Another adaptation has to be done if the channel (\mathcal{R}) has changed over time.

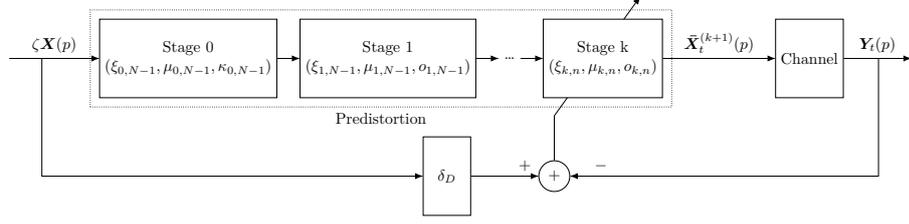
The k -th predistortion output is computed once the parameter adaptation process has converged or when the iteration number n has reached its maximum value (fixed to N).

Once the adaptation process of the k -th stage is over, the gains $(\kappa_n, o_{k,n}, \mu_{k,n}, \xi_{k,n})$ are definitively fixed to their convergence state values $(\kappa, o_k, \mu_k, \xi_k) = (\kappa_N, o_{k,N-1}, \mu_{k,N-1}, \xi_{k,N-1})$.

The stages are adapted one after the other until the pre-defined number of stages K is reached.

Let us denote by $\mathbf{Y}_t(p)$ the p -th nonlinear channel output ($p = kN + n$) vector of length M during the training. Taking into account the delay D introduced by the channel, the predistorted signal $\bar{\mathbf{X}}(p - D)$ is computed so as to minimize the following cost function:

$$J(p) = \mathbf{E}_t^\dagger(p) \mathbf{E}_t(p) \quad (4.18)$$

Figure 4.11: Adaptation of the k -th stage.

where the error vector $\mathbf{E}_t(p)$ is equal to $\zeta \mathbf{X}(p - D) - \mathbf{Y}_t(p)$.

During the k -th stage of the adaptation process, the output of the channel model at the p -th iteration is expressed as:

$$\mathbf{Y}_t(p) = \mathcal{R}[\bar{\mathbf{X}}_t^{(k+1)}(p)], \quad (4.19)$$

where the training predistorted signal $\bar{\mathbf{X}}_t^{(k+1)}(p)$ at iteration p is computed as follows:

$$\bar{\mathbf{X}}_t^{(k+1)}(p) = \begin{cases} \mu_{0,n} \mathbf{X}(p) - \xi_{0,n} \mathcal{R}\{\kappa_n \zeta \mathbf{X}(p)\}, & k = 0 \\ \mu_{k,n} \bar{\mathbf{X}}_t^{(k)}(p) + o_{k,n} \zeta \mathbf{X}(p) - \xi_{k,n} \mathcal{R}[\bar{\mathbf{X}}_t^{(k)}(p)], & k \geq 1 \end{cases} \quad (4.20)$$

with $\bar{\mathbf{X}}_t^{(k)}(p)$ computed by (4.12) or (4.14) with the gains fixed to $(\kappa, o_k, \mu_k, \xi_k)$ obtained after the adaptation process of k -th stage.

Figure 4.11 illustrates the process. In the latter, δ_D denotes the Dirac function delayed by D , which is the delay induced by the channel.

The gains are initialized as follows:

- $\kappa_0 = 1$
- $\mu_{k,0} = 1, \forall k$
- $\xi_{k,0} = 0, \forall k$
- $o_{k,0} = 0, \forall k \geq 1$

We suppose that all the gains are real. Given the k -th predistortion stage output $\bar{\mathbf{X}}_t^{(k+1)}(p)$, the gradient of the cost function for the feedforward parameter μ_k is:

$$\frac{\partial J(p)}{\partial \mu_{k,n}} = -2\Re \left(\mathbf{E}_t^\dagger(p) \left[\frac{\partial \mathbf{Y}_t(p)}{\partial \bar{\mathbf{X}}_t^{(k+1)}(p-D)} \bar{\mathbf{X}}_t^{(k)}(p-D) + \frac{\partial \mathbf{Y}_t(p)}{\partial (\bar{\mathbf{X}}_t^{(k+1)}(p-D))^*} (\bar{\mathbf{X}}_t^{(k)}(p-D))^* \right] \right). \quad (4.21)$$

Then, the feedforward parameter is updated by applying the LMS algorithm, as follows:

$$\mu_{k,n+1} = \mu_{k,n} - c_\mu \frac{\partial J(p)}{\partial \mu_{k,n}}, \quad (4.22)$$

with c_μ a step-size. The other gains $o_{k,n}$, $\xi_{k,n}$ and κ_n are updated in a similar way with the following gradients:

$$\begin{aligned} \frac{\partial J(p)}{\partial o_{k,n}} = & -2\Re \left(\mathbf{E}_t^\dagger(p) \left[\frac{\partial \mathbf{Y}_t(p)}{\partial \bar{\mathbf{X}}^{(k+1)}(p-D)} \mathbf{X}(p-D) \right. \right. \\ & \left. \left. + \frac{\partial \mathbf{Y}_t(p)}{\partial (\bar{\mathbf{X}}_t^{(k+1)}(p-D))^*} (\mathbf{X}(p-D))^* \right] \right), \end{aligned} \quad (4.23)$$

$$\begin{aligned} \frac{\partial J(p)}{\partial \xi_{k,n}} = & 2\Re \left(\mathbf{E}_t^\dagger(p) \left[\frac{\partial \mathbf{Y}_t(p)}{\partial \bar{\mathbf{X}}_t^{(k+1)}(p-D)} \mathcal{R}\{\bar{\mathbf{X}}^{(k)}(p-D)\} \right. \right. \\ & \left. \left. + \frac{\partial \mathbf{Y}_t(p)}{\partial (\bar{\mathbf{X}}_t^{(k+1)}(p-D))^*} (\mathcal{R}\{\bar{\mathbf{X}}^{(k)}(p-D)\})^* \right] \right), \end{aligned} \quad (4.24)$$

$$\begin{aligned} \frac{\partial J(p)}{\partial \kappa_0} = & 2\Re \left(\xi_0 \mathbf{E}_t^\dagger(p) \left[\frac{\partial \mathbf{Y}_t(p)}{\partial \bar{\mathbf{X}}^{(1)}(p-D)} \hat{\mathbf{X}}(p-D) \right. \right. \\ & \left. \left. + \frac{\partial \mathbf{Y}_t(p)}{\partial (\bar{\mathbf{X}}^{(1)}(p-D))^*} (\hat{\mathbf{X}}(p-D))^* \right] \right). \end{aligned} \quad (4.25)$$

with:

$$\hat{\mathbf{X}}(p) = \left(\frac{\partial \mathcal{R}}{\partial \mathbf{Z}} (\kappa_0 \mathbf{X}(p)) \mathbf{X}(p) + \frac{\partial \mathcal{R}}{\partial \mathbf{Z}^*} (\kappa_0 \mathbf{X}(p)) \mathbf{X}^*(p) \right), \quad (4.26)$$

The derivatives $(\frac{\partial \mathcal{R}}{\partial \mathbf{Z}}$ and $\frac{\partial \mathcal{R}}{\partial \mathbf{Z}^*})$ can be computed using equation (2.13) and for a TWTA described by a Saleh's model, using respectively equation (B.5) and equation (B.6).

AMSE

The quality of the adaptation is measured by the normalized mean square error defined as follows:

$$AMSE = 10 \log \left(E \left[\frac{\|\mathbf{Y}(p) - \zeta \mathbf{X}(p-D)\|^2}{\|\zeta \mathbf{X}(p-D)\|^2} \right] \right) \quad [\text{dB}], \quad (4.27)$$

Influence of the pregain κ_0

In this section, we have a look at the effect of the pregain of the convergence performance. Figure 4.12 shows the effect of the pregain gain on the convergence. For the first stage, we see that the pregain allows to reduce the ANMSE while, without this gain, the adaptation has saturated. However, the gain in performance from one stage to the second is reduced after the adaptation of the second stage.

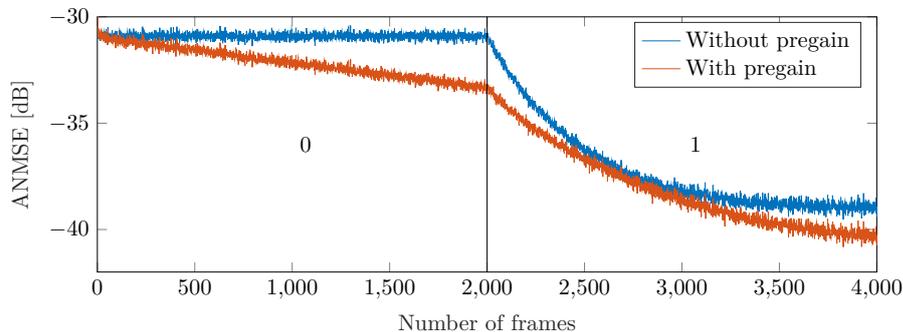


Figure 4.12: ANMSE versus the number of frames with and without pregain.

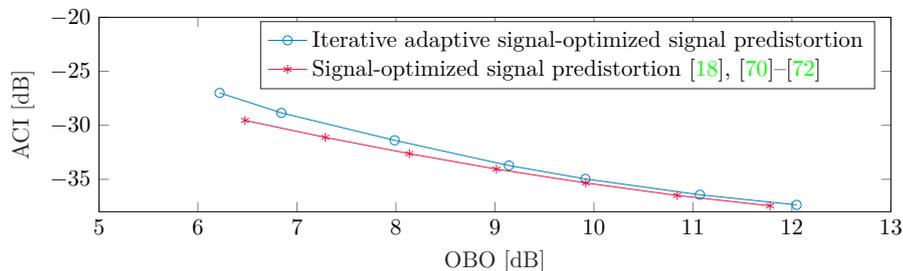


Figure 4.13: ACI at the output of the matched filter versus the OBO.

4.3.4 Comparison with signal-optimized signal predistortion scheme

The simulation setup is the same as in section 4.2.3 except that the amplifier is replaced with a Saleh's model (2,1,2,1.5). The number of symbols per frame N is set to 128 and the oversampling factor to 8 (therefore $M = 1024$).

In this section, we compare the signal-optimized signal predistortion scheme with the proposed iterative adaptive signal predistortion scheme. As the signal-optimized signal predistortion has been optimized for convergence and performance, we can expect a gain in performance.

Results

The ACI factor as a function of OBO is given in Figure 4.13 for both the iterative predistortion based on the fixed point and the proposed iterative adaptive method. We can see that as the OBO increases, the ACI decreases. This phenomenon was illustrated in Section 2.6.3. The reference method outperforms the proposed one by 2.5 dB in the worst case.

On the other hand, concerning the $NMSE$, which relates to the in-band distortion, the trend is reversed. The proposed adaptive method outperforms the reference method, up to 4dB.

The TD with ACI as a function of OBO is plotted in Figure 4.15. The optimal point which minimizes the TD (the lowest along the vertical axis) for the proposed method is better by 1.55dB.

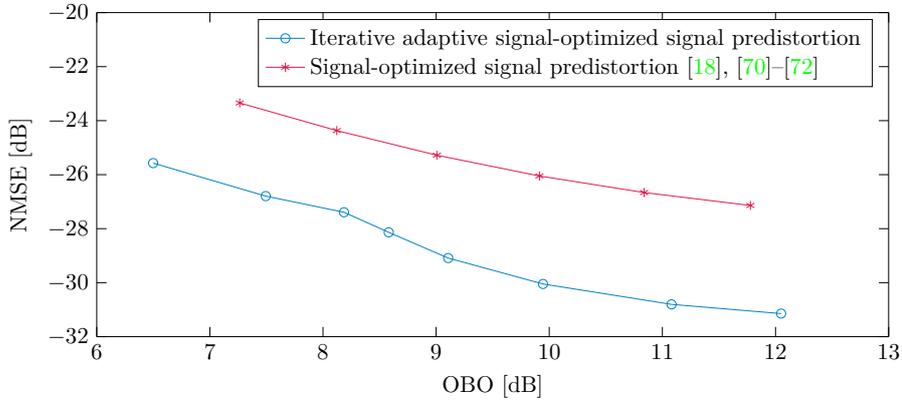


Figure 4.14: NMSE versus the OBO.

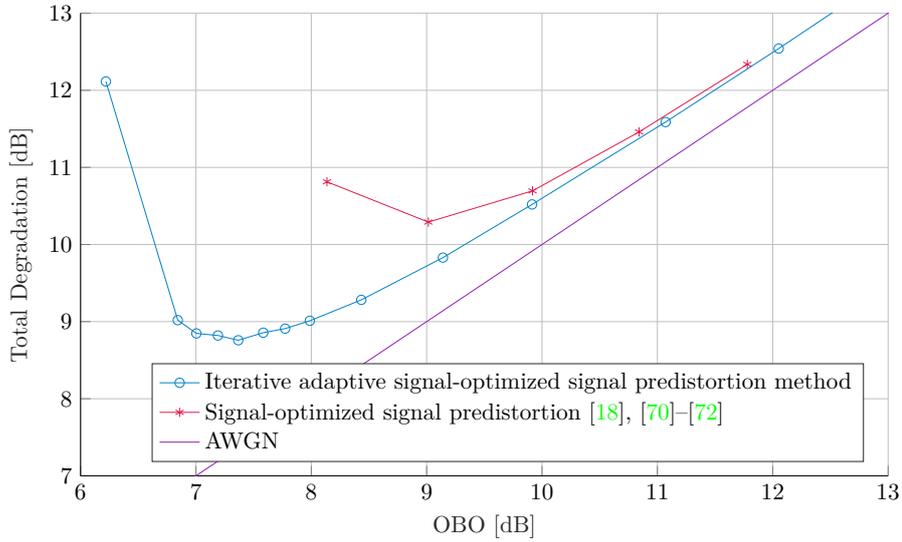


Figure 4.15: Total degradation with ACI for a target BER of 10^{-4} for the reference method and the proposed method for 10 iterations.

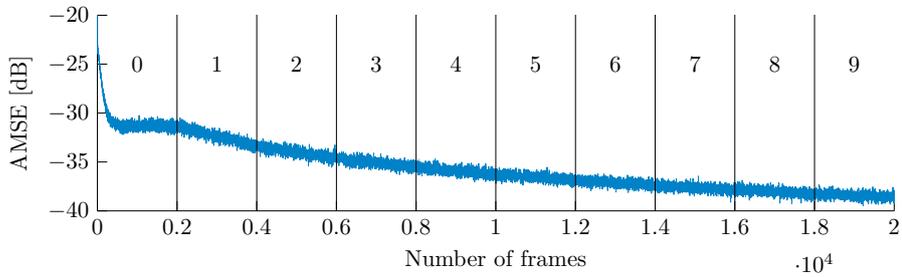


Figure 4.16: AMSE versus number of frames p .

Figure 4.16 illustrates the AMSE defined in (4.27) during the adaptation process as a function of the number of frames p . The upper number mentions the stage under adaptation. When a black vertical line is met, the gains of the current stage are set to the values reached at the end of the adaptive process and the adaptation of the following stage begins. At the first stage output, the adaptation has reached a plateau and the addition of a new stage allows to further decrease the error. The continuity of the AMSE is ensured by the initialization parameter chosen for each stage.

Conclusion

The proposed algorithm circumvents the difficult calculus of convergence conditions and ensures an improvement of linearization whenever the conditions of the contraction mapping theorem are unfulfilled. The complexity of the adaptation process is similar as for a direct learning architecture. Moreover, it avoids empirical tuning of parameters. To do so, three gains were introduced and adapted by a LMS algorithm stage by stage. The issue of exploding or vanishing gradient is avoided by adding the stage one after the other. Monte-Carlo simulations show that, thanks to the gain adaptation, the iterative adaptive signal predistortion outperforms the non-adaptive one in terms of NMSE and TD, with a gain of 1.55 dB for the latter.

4.3.5 Robustness towards feedback mismatch

The simulation setup is the same as in section 4.2.3. Except the amplifier is a Saleh model (see Section 2.4.2) with the following parameter (2,1,2,1).

Equations (4.21) and (4.23) to (4.25) show that the adaptation (and therefore the feedback) is dependent of the derivation of the channel $(\frac{\partial \mathbf{Y}_t(p)}{\partial (\mathbf{X}_t^{(k+1)})^{(p-D)}})^*$, $\frac{\partial \mathbf{Y}_t(p)}{\partial \mathbf{X}^{(k+1)}(p-D)}$. In the following, we want to observe the effect of identification mismatch on the convergence of both data-optimized and signal-optimized signal predistortion. In order to do so, we approximate the nonlinearity of the transponder characteristics using 2.13.

Figure 2.8 shows the approximation of the nonlinear amplifier of the transponder.

Figure 4.17 shows the total degradation for different approximation orders of the feedback for both the signal-optimized and symbol-optimized iterative adaptive signal predistortion. We can see that the approximation of the feedback has a very small impact on the performance of the method.

4.4 Perspectives

Those studies are left for future research works.

4.4.1 Backpropagation

Another way to adapt all the coefficients for a fixed number of stages would be to backpropagate the error through all the steps. The adaptation of the parameters at the same time might improve the performance.

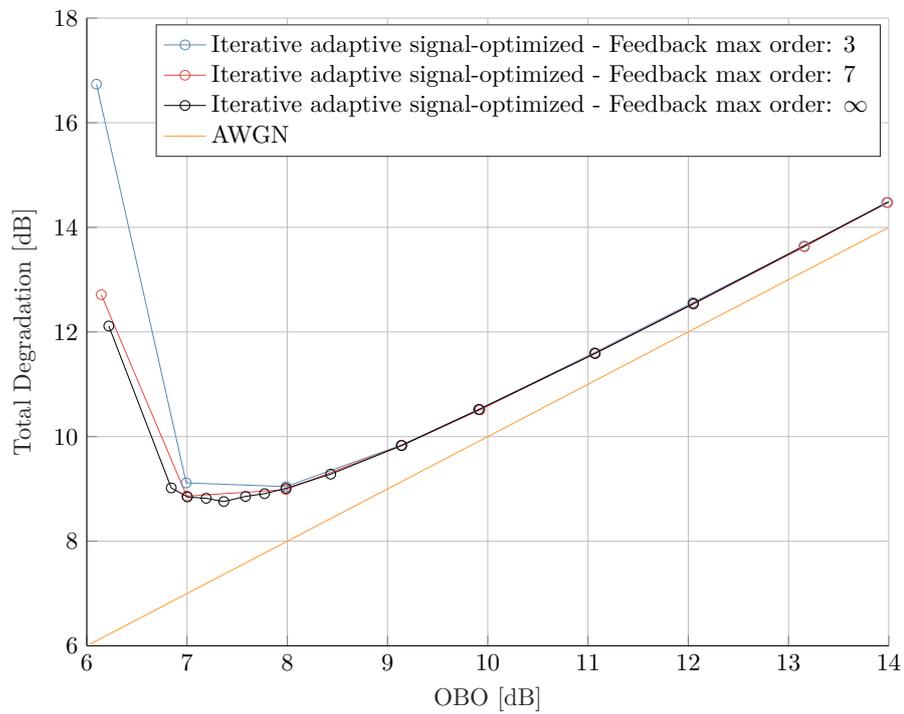


Figure 4.17: Total degradation with ACI for a target BER of 10^{-4} for different feedback approximation.

Let us derive the gradient for the adaptation of the $\mu_{i,n}$:

$$\begin{aligned} \frac{\partial J(p)}{\partial \mu_{i,n}} = & -2\Re \left(\mathbf{E}_t^\dagger(p) \left(\frac{\partial \mathbf{Y}_t(p)}{\partial \bar{\mathbf{X}}^{(k+1)}(p)} \left(\frac{\partial \bar{\mathbf{X}}^{(K+1)}(p)}{\partial \bar{\mathbf{X}}^{(i)}(p)} \frac{\partial \bar{\mathbf{X}}^{(i)}(p)}{\partial \mu_{i,n}} + \frac{\partial \bar{\mathbf{X}}^{(K+1)}(p)}{\partial \bar{\mathbf{X}}^{*(i)}(p)} \frac{\partial \bar{\mathbf{X}}^{*(i)}(p)}{\partial \mu_{i,n}} \right) \right. \right. \\ & \left. \left. + \frac{\partial \mathbf{Y}_t(p)}{\partial \bar{\mathbf{X}}^{*(k+1)}(p)} \left(\frac{\partial \bar{\mathbf{X}}^{*(K+1)}(p)}{\partial \bar{\mathbf{X}}^{(i)}(p)} \frac{\partial \bar{\mathbf{X}}^{(i)}(p)}{\partial \mu_{i,n}} + \frac{\partial \bar{\mathbf{X}}^{*(K+1)}(p)}{\partial \bar{\mathbf{X}}^{*(i)}(p)} \frac{\partial \bar{\mathbf{X}}^{*(i)}(p)}{\partial \mu_{i,n}} \right) \right) \right) \end{aligned} \quad (4.28)$$

Similar equation can be derived for the other parameters.

Let us denote derivatives of the k -th stage linking the output $\bar{\mathbf{X}}_t^{(k+1)}$ and the input $\bar{\mathbf{X}}_t^{(k)}$ by:

$$\frac{\partial \bar{\mathbf{X}}_t^{(k+1)}}{\partial \bar{\mathbf{X}}_t^{(k)}} = \mu_{k,n} I_{N \times N} - \xi_{k,n} \frac{\partial \mathcal{R}[\bar{\mathbf{X}}^{(k)}(p)]}{\partial \bar{\mathbf{X}}^{(k)}(p)} \quad (4.29)$$

$$= \delta_{(k+1,k)}^{(n,l)} \quad (4.30)$$

and

$$\frac{\partial \bar{\mathbf{X}}_t^{(k+1)}}{\partial \bar{\mathbf{X}}_t^{*(k)}} = -\xi_{k,n} \frac{\partial \mathcal{R}[\bar{\mathbf{X}}^{(k)}(p)]}{\partial \bar{\mathbf{X}}^{*(k)}(p)} \quad (4.31)$$

$$= \Delta_{(k+1,k)}^{(n,l)} \quad (4.32)$$

Using the chain rule, the derivative for the previous stage can be computed iteratively:

$$\delta_{(k+1,k-1)}^{(n,l)} = \delta_{(k+1,k)}^{(n,l)} \delta_{(k,k-1)}^{(n,l)} + \Delta_{(k+1,k)}^{(n,l)} \Delta_{(k,k-1)}^{*(n,l)}, \quad (4.33)$$

and

$$\Delta_{(k+1,k-1)}^{(n,l)} = \Delta_{(k+1,k)}^{*(n,l)} \delta_{(k,k-1)}^{(n,l)} + \Delta_{(k+1,k)}^{(n,l)} \delta_{(k,k-1)}^{*(n,l)}. \quad (4.34)$$

Assuming, that the derivative $\Delta_{(k+1,k)}^{(n,l)}$ are all equal to zeros, then,

$$\delta_{(k+1,i)}^{(n,l)} = \prod_{u=i}^k \delta_{(u+1,u)}^{(n,l)} \quad (4.35)$$

$$= \prod_{u=i}^k \left(\mu_{u,n} I_{N \times N} - \xi_{u,n} \frac{\partial \mathcal{R}[\bar{\mathbf{X}}^{(u)}(p)]}{\partial \bar{\mathbf{X}}^{(u)}(p)} \right) \quad (4.36)$$

We can see that the gradient from the first stages are dependent on the value of feedforward $\xi_{k,n}$ and post model gains $\mu_{k,n}$ of all following stages. Then, the adaptation of the parameters of the first stages are prone to suffer from exploding or vanishing gradient.

Thus, the backpropagation might lead to better performance, the complexity cost due to the adaptation is larger and the learning rate has to circumvent this issue of exploding and vanishing gradient.

In comparison to the adaptation proposed in 4.3, the backpropagation is more complex since it has to compute the derivative (equation (4.28)) through the chain rule (equation (4.33) and equation (4.34)).

4.4.2 Joint Clipping and predistortion

The PAPR might be an issue especially for multi-carrier system. In order to circumvent this issue, a solution might be to introduce crest factor reduction in addition to predistortion.

In that context, joint clipping and predistortion have been proposed with Generalized DLA in [15]. This leads to improvement in terms of NMSE compared to the same method without crest factor reduction.

Moreover, in [60], the authors use neural network with a sigmoid activation function. As the output of the structure is the sum of bounded activation functions, the PAPR is reduced with respect to Volterra-derived structure. This also leads to an improvement in terms of NMSE. Therefore, the reduction of the PAPR can be interesting to investigate to improve the performance of the predistortion.

In the case of the signal-optimized or the symbol-optimized signal predistortion, the adaptation of the threshold is deeply intertwined to the nonlinear iterative structure. The reference signal, which has to be clipped, is used at each stage of the predistortion scheme.

Let us now derive the adaptation of the clipping threshold. First, the clipped signal x_c is related to the input by:

$$\frac{\partial x_c(n)}{\partial \gamma(n)} = \begin{cases} 0, & \text{if } |x_c(n)| \leq \gamma, \\ e^{j\angle x_c(n)}, & \text{if } |x_c(n)| > \gamma. \end{cases} \quad (4.37)$$

where γ is the threshold. The derivative is not defined when $|x_c(n)| = \gamma$. In that case, we set the derivative to zero.

Then, the gradient is expressed as follows:

$$\frac{\partial J(p)}{\partial \gamma(p)} = 2\Re \left(\mathbf{E}_t(p) \left(\frac{\partial Y(p)}{\partial \bar{\mathbf{X}}^{(k+1)}(p)} \frac{\partial \bar{\mathbf{X}}^{(k+1)}(p)}{\partial \gamma(p)} + \frac{\partial Y(p)}{\partial \bar{\mathbf{X}}^{*(k+1)}(p)} \frac{\partial \bar{\mathbf{X}}^{*(k+1)}(p)}{\partial \gamma(p)} \right) \right), \quad (4.38)$$

with

$$\begin{aligned} \frac{\partial \bar{\mathbf{X}}^{(k+1)}(p)}{\partial \gamma(p)} &= o_k \zeta \frac{\partial \mathbf{X}_c(p)}{\partial \gamma(p)} + \mu_k \frac{\partial \bar{\mathbf{X}}^{(k)}(p)}{\partial \gamma(p)} \\ &\quad - \xi_k \left(\frac{\partial \mathcal{R}[\bar{\mathbf{X}}^{(k)}(p)]}{\partial \bar{\mathbf{X}}^{(k)}(p)} \frac{\partial \bar{\mathbf{X}}^{(k)}(p)}{\partial \gamma(p)} + \frac{\partial \mathcal{R}[\bar{\mathbf{X}}^{(k)}(p)]}{\partial \bar{\mathbf{X}}^{*(k)}(p)} \frac{\partial \bar{\mathbf{X}}^{*(k)}(p)}{\partial \gamma(p)} \right), \end{aligned} \quad (4.39)$$

and

$$\frac{\partial \bar{\mathbf{X}}^{(1)}(p)}{\partial \gamma(p)} = \mu_0 \frac{\partial \mathbf{X}_c(p)}{\partial \gamma(p)} - \xi_0 \left(\frac{\partial \mathcal{R}[\mathbf{X}_c(p)]}{\partial \mathbf{X}_c(p)} \frac{\partial \mathbf{X}_c(p)}{\partial \gamma(p)} + \frac{\partial \mathcal{R}[\mathbf{X}_c(p)]}{\partial \mathbf{X}_c^*(p)} \frac{\partial \mathbf{X}_c^*(p)}{\partial \gamma(p)} \right). \quad (4.40)$$

Like for the backpropagation, the gradient is likely to explode.

4.4.3 Variable step-size

Using variable step-size might be a solution to circumvent the issue of vanishing or exploding gradient. For instance, for the backpropagation, it validates most

of the heuristics detailed in [64]. This solution has already been proposed with DLA in [78] for narrowband signal.

The variable allows to adapt the learning rate at every iteration. It is respectively increased or decreased, if for several consecutive iterations, the gradient has the same algebraic sign.

Let us now derive the adaptation of the step-size for the post model gain. The step-size $\alpha_{\xi,n}$ at the n -th iteration of the parameter ξ reads:

$$\alpha_{\xi,n} = \alpha_{\xi,n-1} - \beta_{\xi} \frac{\partial J(n)}{\partial \alpha_{\xi,n-1}} \quad (4.41)$$

with β_{ξ} is a constant step-size.

The cost function is:

$$J(n) = \|\zeta \mathbf{X}(n - D) - \mathbf{Y}(n)\|^2, \quad (4.42)$$

The parameter ξ_k is updated as follows:

$$\xi_{k,n} = \xi_{k,n-1} - \alpha_{\xi,n-1} \frac{\partial J(n)}{\partial \xi_{k,n-1}}. \quad (4.43)$$

Then,

$$\frac{\partial J(n)}{\partial \alpha_{\xi,n-1}} = -E^{\dagger}(l) \frac{\partial Y(l)}{\partial \alpha_{\xi,n-1}} - E(l) \frac{\partial Y^*(l)}{\partial \alpha_{\xi,n-1}} \quad (4.44)$$

$$= -2\Re(E^{\dagger}(l) \frac{\partial Y(l)}{\partial \alpha_{\xi,n-1}}) \quad (4.45)$$

with

$$\frac{\partial Y(l)}{\partial \alpha_{\xi,n-1}} = \left(\frac{\partial Y(l)}{\partial \bar{\mathbf{X}}^{(k)}(n)} \frac{\partial \bar{\mathbf{X}}^{(k)}(n)}{\partial \xi_{k,n}} + \frac{\partial Y(l)}{\partial \bar{\mathbf{X}}^{*(k)}(n)} \frac{\partial \bar{\mathbf{X}}^{*(k)}(n)}{\partial \xi_{k,n}} \right) \frac{\partial \xi_{k,n}}{\partial \alpha_{\xi,n-1}} \quad (4.46)$$

Therefore, we have:

$$\frac{\partial J(n)}{\partial \alpha_{\xi,n-1}} = -2\Re(E^{\dagger}(n) \frac{\partial Y(n)}{\partial \alpha_{\xi,n-1}}) \quad (4.47)$$

$$= -\frac{\partial J(n)}{\partial \xi_{k,n}} \frac{\partial J(n-1)}{\partial \xi_{k,n-1}} \quad (4.48)$$

The variable step-sizes for the other parameters (μ_k , σ_k and κ_0) can be derived in a similar fashion.

4.5 Conclusion

In this chapter, we try to address the problem of predistortion by proposing a data predistortion based on the contraction mapping. This scheme obtains good performance with respect to the state-of-the-art method. However, the amplifier working point that minimizes the TD might not satisfy the conditions of the contraction mapping theorem.

Therefore, we propose to introduce parameters in order to ensure the convergence and the best performance for a fixed number of iterations. The proposed schemes are, then, twice adaptive.

The proposed method adapts the introduced coefficient by back-propagating the gradient through the channel which avoids empirical tuning of parameter.

This scheme is resilient to feedback mismatch. The resilience toward feedback mismatch might be due to the fact that the structure is trained based on [DLA](#).

To sum up, the proposed methods only require an identification of the channel, it is needless to set a prior on the structure of the predistortion and to fine tune parameters.

Chapter 5

Conclusion and perspectives

“What’s gonna to happen ? Probably a fistfight.”

Jack White as himself.
It Might Get Loud. Directed by Davis Guggenheim.
Steel Curtain Pictures. 2008

Contents

5.1 Conclusion	83
5.2 Perspectives	84
5.2.1 Similarity with Neural Network	84
5.2.2 Use of piecewise-defined function	85
5.2.3 Iterative processing for FTN	85
5.2.4 Joint adaptation of the constellation and predistortion	85

5.1 Conclusion

In this thesis, we focus on the management of distortion for satellite communications. First, we study the clustering and wrapping effect for multiple rings constellation with evenly distributed symbols (on each ring). We show that the conditional variance of the received symbol depends on the amplitude of the rings and may also depend on the phase of the transmitted signal (according to the modulation). Therefore, given the order of the nonlinearity and the modulation constellation, it may be required to process the real and imaginary part separately.

In a second movement, we study counter-measures for the wrapping and clustering effect at the transmitter. We restrict our study to iterative predistortion since according to the literature, it offers better performance than direct one. The problem of predistortion at symbol rate is first reformulated as a fixed-point problem. When the necessary conditions of the contraction mapping theorem are satisfied, the solution is unique and can be found iteratively. Then, as data

predistortion does not increase the spectral support of the uplink carrier, we propose a data predistortion based on the contraction mapping theorem. In practical case, the channel has to be identified. We compare, in that context, the proposed method with a similar signal-optimized predistortion. Simulations show that the data predistortion is more robust to identification mismatch. However, a damping factor may be necessary to ensure the convergence. In order to circumvent the issue of empirical tuning of parameter, we introduce gains that can be adapted with a feedback from the channel. Those extra gains are used to ensure the best performance after a given number of iterations as well as convergence when the conditions of the contraction mapping theorem are unfulfilled. We propose to adapt the additional coefficients with a [DLA](#), as the nonlinear processing from stage to stage might increase the noise with an [ILA](#). We design an adaptation scheme that avoids the issue of exploding or vanishing gradient that can be met with backpropagation adaptation method. And lastly, we partially investigate the resilience of the adaptation method towards feedback error.

In a nutshell, the twice-adaptive predistortion scheme allows to build a predistortion without having to give a priori on the structure of the predistorter since the predistorted signal is obtained thanks to the model of the channel. The data-optimized processes are more resilient to identification mismatch, they do not increase the spectral support of the uplink. However, they do not reduce the spectral regrowth on the downlink. The signal-optimized processes can offer better performance, but, they tend to be more sensitive to identification mismatch. The spectral support of the uplink is increased but the spectral regrowth is reduced. Therefore, according to the constraints of the system design, one solution over the other can be favoured.

5.2 Perspectives

In addition to the perspectives already mentioned in [Section 4.4](#), we could investigate the following tracks.

5.2.1 Similarity with Neural Network

In [chapter 4](#), we discarded the back-propagation method to adapt the proposed scheme due to its potential exploding or vanishing gradient. However, as the structure of the proposed solutions have a similar structure as Neural Network, the back-propagation could lead to better performance.

The similarity between the proposed iterative structure and neural network is more noticeable if we use a Wiener-Hammerstein model inside each step. The update equation of both the signal-optimized and data-optimized signal predistortion could be expressed as a Wiener-Hammerstein system. As we argued in [Section 3.2.4](#), the [MLP](#) is similar to Wiener-Hammerstein model. Therefore, could solutions proposed be seen as neural networks ?

As [\[79\]](#), the set of functions that neural network can computed (expressivity) increased exponentially with the depth of the neural network. Could it help to justify theoretically the use of iterative process over direct one ?

5.2.2 Use of piecewise-defined function

We could use a piecewise-defined function instead of a polynomial approximation to model the satellite transponder in order to reduce the complexity of the model.

5.2.3 Iterative processing for FTN

As it was highlighted in [74], the euclidean distance is not reduced for pulse satisfying the Nyquist criterion if the symbol rate is increased until the Mazo's limit is reached. Then, beyond the Mazo's limit, the euclidean distance between two distinct sequences is reduced, which can be seen as a contraction mapping.

It should be possible to equalize FTN signals at the receiver (beyond the Mazo's limits). However, the solution that would be provided by the Contraction Mapping Theorem should lead to zero-forcing equalization, which would enhance the noise. The update equation to recover the transmitted sequence could be:

$$\tilde{\mathbf{d}}^{(k+1)} = (\mathbf{I} - \mathbf{G}_{Rx}\mathbf{H}_{Tx})\tilde{\mathbf{d}}^{(k)} + \hat{\mathbf{d}} \quad (5.1)$$

where $\tilde{\mathbf{d}}$ is the reconstructed transmitted sequence and $\hat{\mathbf{d}}$ is the received sequence.

Likewise, similar iterative process can be used to recover the transmitted sequences when time-frequency packing over the two-dimensional Mazo's limits [80].

5.2.4 Joint adaptation of the constellation and predistortion

For a predefined modulation order, the joint adaptation of the predistortion and the constellation could be investigated. The receiver could track the modification of the constellation and compensate part of the wrapping effect by using a centroid-based decoder. The constellation and the working point of the amplifier could be optimized in a similar fashion to the joint clipping and predistortion proposed in Section 4.4. The method could favour Phase-shift keying (PSK) constellation for amplifier with small dynamic (which heavily compressed the signal) and APSK could be used otherwise. One issue with that approach is that the bit mapping to the optimized constellation could be sub-optimal, and therefore, it could thwart the benefit of the optimisation.

Appendix A

Derivation of the conditional variance and pseudo-variance for a third order Volterra model

Contents

A.1 Mathematical foreword	88
A.2 Conditional variance for a third order Volterra model	88
A.3 Conditional pseudo-variance for a third order Volterra model	89

Let us now describe the clustering effect. The clustering effect can be seen as an additional noise due to the distortion.

As we work with complex constellation, we need to both to derive the variance and the pseudo-variance in order to characterize the clustering effect.

The derivation is valid for a large of constellation that can be written as:

$$\{\{a_i e^{2j\pi \frac{k}{n_i}}, i = [1, R], a_i \in \mathbb{C}, \forall n_i = 2l, l \in \mathbb{N}^*, k = [0, n_i - 1]\}, \forall R \in \mathbb{N}^*\} \quad (\text{A.1})$$

where R is the number of rings, n_i (≥ 2) is the even number of symbols per rings, a_i is the amplitude and phase shift of the i -th ring.

Therefore, for any symbol randomly drawn from the constellation:

$$\mathbb{E}[d_n] = 0 \quad (\text{A.2})$$

As those constellations consist of stacked scaled and phase shifted of root of unity, we have:

$$\mathbb{E}[d_n^{2m+1}] = 0 \quad (\text{A.3})$$

for any order m (≥ 1) and

$$\mathbb{E}[d_n^{2m}] = \sum_{i=1}^R a_i^{2m} \delta_{((2m \bmod n_i), 0)} \quad (\text{A.4})$$

where $\delta_{(i,j)}$ is the Kronecker delta defines as follows:

$$\delta_{(i,j)} = \begin{cases} 1, & \text{for } i = j, \\ 0, & \text{otherwise} \end{cases} \quad (\text{A.5})$$

A.1 Mathematical foreword

Let us define:

$$d_n = u_n + jv_n \quad (\text{A.6})$$

where u_n and v_n are respectively, the real part and the imaginary part of d_n .

The variance is defined as:

$$\mathbb{E}[d_n d_n^*] = R_{u_n u_n} + R_{v_n v_n}, \quad (\text{A.7})$$

where $R_{u_n u_n}$ is the $R_{v_n v_n}$ are respectively, the variance of the real part and the imaginary part, and the pseudo-variance as:

$$\mathbb{E}[d_n d_n] = R_{u_n u_n} - R_{v_n v_n} - 2j\sqrt{(R_{u_n u_n} R_{v_n v_n})}\rho_{u_n v_n}. \quad (\text{A.8})$$

where $\rho_{u_n v_n}$ is the correlation between the real u_n and the imaginary part v_n .

If the imaginary part of the (conditional) pseudo-variance is non zero, then the real and imaginary part of the received symbol are correlated. Likewise, if the real part of the (conditional) pseudo-variance is non zero, then the variance of the real and imaginary part of the received symbol are not equal. Therefore, this also imply that the distribution is elliptic [81].

A.2 Conditional variance for a third order Volterra model

First let us derive the conditional variance for a third order nonlinearity and using equation (2.19), the conditional variance reads:

$$\mathbb{E}[|\hat{d}_n - \mathbb{E}[\hat{d}_n]|^2 | d_0 = d] = \mathbb{E}[|\hat{d}_n|^2 | d_n = d] - |\mathbb{E}[\hat{d}_n | d_n = d]|^2 \quad (\text{A.9})$$

with

$$\begin{aligned} \mathbb{E}[|\hat{d}_n|^2 | d_n = d] &= \sum_{l,l'} q_{(1)l} q_{(1)l'}^* \mathbb{E}[d_{n-l} d_{n-l'}^* | d_n = d] \\ &+ \sum_{i,j,k',l'} q_{(3)i,j,k'} q_{(1)l'}^* \mathbb{E}[d_{n-i} d_{n-j} d_{n-k'}^* d_{n-l'}^* | d_n = d] \\ &+ \sum_{i',j',k,l} q_{(3)i',j',k} q_{(1)l} \mathbb{E}[d_{n-i'}^* d_{n-j'}^* d_{n-k} d_{n-l} | d_n = d] \\ &+ \sum_{i,j,k,i',j',k'} q_{(3)i,j,k} q_{(3)i',j',k'}^* \mathbb{E}[d_{n-i} d_{n-j} d_{n-k}^* d_{n-i'}^* d_{n-j'}^* d_{n-k'} | d_n = d] \end{aligned} \quad (\text{A.10})$$

$$\mathbb{E}[d_l d_{l'}^* | d_0 = d] = \begin{cases} |d|^2 \\ \sigma_c^2 \\ 0 \end{cases} \quad \text{otherwise} \quad (\text{A.11})$$

$$\mathbb{E}[d_i d_j d_k^* d_l^* | d_0 = d] = \begin{cases} |d|^4 \\ \sigma_c^2 |d|^2 \\ R_4^{2,2} \\ \sigma_c^4 \\ |R_2^{2,0}|^2 \\ R_2^{2,0} (d^*)^2 \\ R_2^{0,2} d^2 \\ 0 \end{cases} \quad \text{otherwise} \quad (\text{A.12})$$

$$\mathbb{E}[d_i d_j d_k d_l^* d_m^* d_n^* | d_0 = d] = \begin{cases} |d|^6 \\ R_6^{3,3} \\ R_4^{2,2} |d|^2 \\ |R_2^{2,0}|^2 |d|^2 \\ \sigma_c^4 |d|^2 \\ \sigma_c^2 |d|^4 \\ \sigma_c^6 \\ \sigma_c^2 R_2^{0,2} d^2 \\ \sigma_c^2 |R_2^{2,0}|^2 \\ 0 \end{cases} \quad \text{otherwise} \quad (\text{A.13})$$

We see that the conditional variance of the received symbol depends on the statistics of the chosen constellation (σ_c^2, R_4, R_6) and the modulus of the symbol under consideration ($|d|$). Therefore, the symbols on the outer rings of multiple rings constellation tend to be more affected by the distortion than the one on the inner rings since the modulus $|d|$ is greater on the outer rings.

A.3 Conditional pseudo-variance for a third order Volterra model

Let now compute the pseudo-covariance:

$$\mathbb{E}[(\hat{d}_n - \mathbb{E}[\hat{d}_n])^2 | d_0 = d] = \mathbb{E}[(\hat{d}_n)^2 | d_n = d] - \mathbb{E}[\hat{d}_n | d_n = d]^2 \quad (\text{A.14})$$

with

$$\begin{aligned} \mathbb{E}[\hat{d}_n^2 | d_n = d] &= \sum_{l, l'} q_{(1)l} q_{(1)l'} \mathbb{E}[d_{n-l} d_{n-l'} | d_n = d] \\ &+ 2\Re \left(\sum_{i, j, k, l} q_{(3)i, j, k} q_{(1)l} \mathbb{E}[d_{n-i} d_{n-j} d_{n-k}^* d_{n-l'} | d_n = d] \right) \\ &+ \sum_{i, j, k, i', j', k'} q_{(3)i, j, k} q_{(3)i', j', k'} \mathbb{E}[d_{n-i} d_{n-j} d_{n-k}^* d_{n-i'} d_{n-j'} d_{n-k'}^* | d_n = d] \end{aligned} \quad (\text{A.15})$$

$$\mathbb{E}[d_l d_{l'} | d_0 = d] = \begin{cases} d^2 & (l = l' = 0) \\ R_2^{2,0} & (l = l' \neq 0) \\ 0 & \text{otherwise} \end{cases} \quad (\text{A.16})$$

$$\mathbb{E}[d_i d_j d_k^* d_l | d_0 = d] = \begin{cases} |d|^2 d^2 & \text{for } i = j = k' = l = 0 \\ |d|^2 R_2^{2,0} & \\ \sigma_c^2 d^2 & \\ 0 & \text{otherwise} \end{cases} \quad (\text{A.17})$$

$$\mathbb{E}[d_i d_j d_k^* d_{i'} d_{j'} d_{k'}^* | d_0 = d] = \begin{cases} R_4^{2,2} d^2 \\ |R_2^{2,0}|^2 d^2 \\ \sigma_c^4 d^2 \\ |d|^4 d^2 \\ \sigma_c^2 |d|^2 d^2 \\ R_4^{2,2} d^2 \\ |R_2^{2,0}|^2 d^2 \\ \sigma_c^4 R_2^{2,0} \\ |d|^4 R_2^{2,0} \\ \sigma_c^2 |d|^2 R_2^{2,0} \\ R_4^{4,0} R_2^{0,2} \\ 0 & \text{otherwise} \end{cases} \quad (\text{A.18})$$

Assuming that the filter have real coefficient, then, those properties will depend on the product of coefficients of Taylor expansion of the characteristics of the amplifier γ_{2m+1} and d^2 . Therefore, in most cases, it can be expect that the conditional distribution to be elliptic.

Similar result can be easily derived for higher order Volterra model based on the previous derivation. In the case of higher order Volterra model, higher order statistics of the transmitted constellation are required.

Appendix B

Computation of the derivative of Saleh's HPA model

Contents

B.1 Derivative of Saleh's HPA model	91
---	-----------

B.1 Derivative of Saleh's HPA model

A commonly used model for [TWTA](#) is described in [\[28\]](#). The Saleh model uses four parameters $(\alpha_a, \beta_a, \alpha_\phi, \beta_\phi)$:

$$A(\rho(t)) = \alpha_a \frac{\rho(t)}{1 + \beta_a \rho^2(t)}, \quad (\text{B.1})$$

$$\Omega(\rho(t)) = \alpha_\phi \frac{\rho^2(t)}{1 + \beta_\phi \rho^2(t)}, \quad (\text{B.2})$$

with $\rho^2(t)$ is the instantaneous signal power. This model can be made frequency-dependent by adding a filter [\[28\]](#), [\[29\]](#).

The Saleh model is a four-parameters $(\alpha_A, \beta_A, \alpha_\Omega, \beta_\Omega)$ model. The [AM-AM](#) characteristics is given by:

$$A(\rho(t)) = \alpha_A \frac{\rho(t)}{1 + \beta_A \rho(t)^2}, \quad (\text{B.3})$$

and the [AM-PM](#) is given by:

$$\Omega(\rho(t)) = \alpha_\Omega \frac{\rho(t)^2}{1 + \beta_\Omega \rho(t)^2}, \quad (\text{B.4})$$

with $\rho = \sqrt{xx^*}$ is the square root of the [Instantaneous Normalized Power \(INP\)](#).

Let us denote the input signal of the Saleh model as: $x(t) = \rho(t) \exp(j\angle x(t))$ and the output is, then: $y(t) = A(\rho(t)) \exp(j[\angle x(t) + \Omega(\rho(t))])$

The derivatives reads:

$$\frac{\partial y}{\partial x} = \alpha_A \exp(j\Omega(\rho)) \left(\frac{(1 + \beta_\Omega \rho^2)^2 + j\alpha_\Omega \rho^2 (1 + \beta_A \rho^2)}{(1 + \beta_A \rho^2)^2 (1 + \beta_\Omega \rho^2)^2} \right), \quad (\text{B.5})$$

and

$$\frac{\partial y}{\partial x^*} = \alpha_A \exp(j\Omega(\rho)) x^2 \left(\frac{j\alpha_\Omega (1 + \beta_A \rho^2) - \beta_A (1 + \beta_\Omega \rho^2)^2}{(1 + \beta_A \rho^2)^2 (1 + \beta_\Omega \rho^2)^2} \right). \quad (\text{B.6})$$

Appendix C

Résumé

C.1 Introduction et contexte

Les communications satellitaires sont apparues après la fin de la Seconde Guerre Mondiale. Le développement des missiles à longues portées durant la guerre, et notamment les *V2*, par les Allemands, a permis le début de la conquête spatiale et, par conséquent le lancement de satellites artificiels.

Le principal avantage d'utiliser des relais extra-terrestres sur n'importe quel système de communications terrestres est la capacité de couvrir de larges zones. A tel point qu'il ne suffit que de trois satellites géostationnaires pour assurer une couverture mondiale. Les liens de communications satellitaires sont donc compétitifs pour les applications de radiodiffusion.

Un autre avantage des communications par satellite, est la résilience vis-à-vis des désastres technologiques et naturels. Par exemple, le programme de l'agence spatiale européenne, EMERGSAT, terminé en 2008, avait pour but de fournir des liens de communications entre plusieurs stations fixes ou mobiles dans les heures après une catastrophe en Europe et dans le bassin méditerranéen.

Par la suite, le déploiement de lourdes infrastructures terrestres pour connecter le réseau de backhaul des réseaux de téléphonie mobile n'est pas nécessaire pour couvrir les zones mal ou non desservies est aussi un avantage pour le déploiement de réseaux satellitaires dans les pays en développement.

Avec la réduction des coûts de lancement des satellites, grâce à la réutilisation partielle et l'accroissement des capacités des lanceurs, des grandes constellations de petits satellites sont en cours de développement ou en cours de lancement par plusieurs compagnies privées (ces constellations se situent en orbite basse ou moyenne) avec pour objectifs de procurer une connectivité internet à faible latence.

Pour les services Wi-fi maritimes, aéroportés et ferroviaires à grandes vitesses, la solution satellitaire semble être une bonne alternative aux solutions terrestres. Il est facile de couvrir de large distance et les grandes étendues d'eau, les terrains isolés et zones rurales au travers de liens de communications satellitaires.

Ces avantages, la possibilité de fournir une couverture radio partout et plus simplement pour les applications maritimes, aéroportés et ferroviaires et le potentiel pour connecter le réseau de backhaul ce qui facilite la modularité du réseau de la *cinquième génération de téléphonie mobile* (5G): poussent l'inté-

gration d'un volet satellitaire dans la 5G.

Toutes les applications susmentionnées nécessitent des débits plus élevés sur les liaisons de communications par satellite, ce qui implique une amélioration technologique des solutions proposées jusqu'alors.

Dans le cas de la diffusion vidéo, l'introduction de la télévision à haute définition et à ultra-haute définition (UHDTV) et des reportages d'actualités par satellite (RAS) constituent la majorité du besoin d'augmentation de débit.

Le DVB-S a été publié pour la première fois en 1994, et était principalement conçu pour la radiodiffusion télévisuelle directe à domicile par satellite. Il n'offre pas beaucoup de flexibilité en terme d'efficacité spectrale puisqu'il était contraint par l'utilisation d'une seule modulation, la MDP-4, et d'un unique facteur de retombé de 35%.

Ce problème a mené au développement du DVB-S2, qui est paru en 2005. Pour fournir plus de capacité, des ordres de modulation (de la MDP-8 à la 32-APSK) et des facteurs de retombés plus marqués (jusqu'à 20%) ont été introduits.

L'extension du standard DVB-S2, le DVB-S2X, a été approuvé en 2014. Il permet d'opérer à très faible et à très fort rapport signal à bruit et augmente la granularité des MODCOD (combinaison de modulation et de codage) disponible. Pour ce faire, le DVB-S2X a introduit, parmi les autres solutions, des ordres de modulations encore plus élevés (de la 64-APSK à la 256-APSK) mais aussi des facteurs de retombées encore plus franc (jusqu'à 5%).

Pour les applications de radiodiffusion et celles à haut débit par satellite, qui sont souvent des applications directes à domicile, une topologie de réseau en étoile est souvent utilisée. C'est une connectivité point à multipoints. Dans ce cas, l'émetteur envoie un signal vers le satellite, puis le satellite amplifie et souvent change la fréquence porteuse du signal reçu pour enfin le retransmettre vers les récepteurs terrestres.

Les systèmes de radiodiffusion par satellite consiste d'un émetteur qui envoie les données à des utilisateurs finaux (récepteurs) par satellite. Le lien de communication entre l'émetteur et le satellite est appelé lien montant tandis que celui entre le satellite et les utilisateurs finaux est appelé lien descendant.

La principale source de dégradation sur ce type de lien de communication est dû au transpondeur satellitaire. La réponse de l'amplificateur embarqué est souvent nonlinéaire, ce qui cause l'apparition d'interférences entre-symboles non-linéaires. La distorsion peut être évitée au coût de l'augmentation du recul en sortie de l'amplificateur (OBO), ce qui réduit la puissance reçue par les récepteurs mais permet à l'amplificateur d'être plus linéaire. Cependant, les contraintes de poids et de puissance imposées par le satellite empêche cette solution.

La distorsion introduite est accrue lorsque l'on utilise des ordres de modulation élevés et des facteurs de retombées plus strictes puisqu'ils augmentent le ratio crête sur moyenne, et donc amène l'amplificateur à saturer plus.

Les figures 1.2, 1.3 et 1.4, montrent respectivement qu'un recul à l'entrée de l'amplificateur plus petit, qu'un ordre de modulation plus élevé et qu'un facteur de retombé plus marqué impliquent une plus forte dégradation introduite par la distorsion.

Comme nous l'avons vu, les méthodes classiques pour augmenter le débit, comme accroître l'ordre de modulation et utiliser un facteur de retombé plus étroit, ont tendance à amplifier l'interférence entre-symboles nonlinéaire. Par

conséquent, pour augmenter le débit, la compensation des nonlinéarités est une technologie clé.

Il y a plusieurs façons d'aborder ce problème, parmi les solutions communément trouvées dans la littérature, l'on peut citer la réduction du facteur de crête, la prédistorsion, la linéarisation embarquée et l'égalisation.

La réduction du facteur de crête est souvent appliquée pour les systèmes multi-porteuses plutôt que les systèmes mono-porteuses car le facteur de crête est moins important avec ces derniers.

Dans un contexte de radiodiffusion, les récepteurs sont sensibles aux coûts. La capacité de calcul doit donc être limitée au minimum. De plus l'égalisation nonlinéaire peut réhausser le niveau de bruit.

La linéarisation embarquée peut être l'une des méthodes la plus efficace puisqu'elle ne doit compenser uniquement la réponse nonlinéaire de l'amplificateur et non la totalité du transpondeur. Par conséquent, l'effet mémoire introduit par le filtrage avant et après la nonlinéarité n'est pas à prendre en compte. Cependant, ces solutions impliquent une augmentation de la charge et donc une augmentation du coût de lancement du satellite.

La prédistorsion est donc souvent une bonne façon de compenser les nonlinéarités, tout en gardant le coût des récepteurs et du satellite bas. En effet, la puissance de calcul disponible à l'émetteur est souvent suffisante pour réaliser l'opération de prédistorsion.

Dans la littérature, la plupart des techniques de prédistorsion, la performance de ces dernières dépendent de la structure choisie du prédistorseur. D'autres méthodes se reposent sur le choix d'une structure pour identifier la réponse du satellite pour laquelle on dispose de plus d'informations a priori.

Les méthodes itératives semblent surpasser les méthodes directes en terme de performance de linéarisation mais celle-ci nécessite souvent un réglage empirique de paramètres.

Cette thèse se concentre sur la compensation des nonlinéarités pour le DVB-S2. Dans un premier mouvement, pour contourner les deux problèmes de définir une structure de prédistorseur a priori et d'adapter empiriquement des paramètres, nous avons proposé une méthode itérative basée sur le théorème du point fixe visant à minimiser l'erreur quadratique moyenne normalisée. La méthode proposée nécessite seulement l'identification du canal. Les performances sont donc liées à la qualité de l'identification et non du choix de la structure.

Cependant, la vérification des hypothèses du théorème du point fixe est dépendant de la norme choisie. De plus, la meilleure performance en terme de dégradation totale est atteinte lorsqu'un compromis entre la compensation des nonlinéarités et le recul à la sortie de l'amplificateur est accomplie. Par conséquent, dans certains cas, des distorsions subsistent et donc ne tombent dans le cadre du théorème. Néanmoins, faire quelques itérations peut améliorer la linéarité de l'amplificateur.

Un autre problème survient lorsque l'on identifie le canal: l'inadéquation entre les structures requière l'introduction de paramètres empiriques. Par conséquent, nous avons introduit des paramètres et proposé une méthode d'adaptation qui assure l'amélioration de la linéarité.

C.2 Communications sur un lien satellitaire

Dans ce chapitre est présenté le contexte de communication par satellite.

La chaîne de communication comprend trois éléments, l'émetteur, le transpondeur du satellite et le récepteur.

L'émetteur et le transpondeur sont décrits dans le standard DVB-S2(X). Nous nous intéressons ici, pour les caractéristiques de l'émetteur aux constellations choisies ainsi que les différents filtres d'émission.

Concernant le transpondeur qui est composé des éléments suivants, le filtre IMUX, qui sélectionne la porteuse d'intérêt sur le lien montant, l'amplificateur, et le filtre OMUX, qui permet de limiter l'accroissement spectral sont présentés.

Les caractéristiques des filtres IMUX et OMUX viennent du standard DVB-S2(X).

Par la suite, les différents modèles d'amplificateur présents dans la littérature sont décrits. En pratique, deux types d'amplificateurs sont principalement utilisés, les tubes à ondes progressives et les amplificateurs de puissance à semi-conducteurs. Les amplificateurs de puissance à semi-conducteurs ayant des caractéristiques plus linéaires que les tubes à ondes progressives, nous allons uniquement nous concentrer sur ces derniers. Ce point est souligné par la présence dans le standard uniquement de caractéristiques pour les tubes à ondes progressives, les amplificateurs de puissance à semi-conducteurs étant jugés moins critiques.

La cascade de filtres linéaires et d'un amplificateur non-linéaire (répartis entre l'émission, le transpondeur et le récepteur) forment un système non-linéaire avec de la mémoire.

Nous décrivons donc les différents modèles utilisés dans la littérature et leurs hypothèses sous-jacentes. Sont présentés, le modèle de Wiener-Hammerstein qui est une représentation par bloc de la chaîne, le modèle de Volterra, qui propose par le biais d'une décomposition de Taylor, une forme analytique permettant de relier les symboles émis et reçus, et enfin, le modèle polynomial à mémoire qui est une version simplifiée du précédent modèle.

Par le biais du modèle de Volterra, l'effet de la distorsion sur la constellation reçue est étudiée statistiquement, pour les constellations constituées de plusieurs anneaux (niveaux d'énergie) ayant un nombre pair de symboles par anneau uniformément répartis. Il est montré que dans la majorité des cas, les symboles reçus ne sont pas invariants par rotation. Ces derniers peuvent être invariants par rotation s'il y a suffisamment de symboles par anneau dans la constellation. Il est donc souvent nécessaire de traiter la partie réelle et imaginaire des symboles reçus séparément.

Ce résultat est illustré par simulation, voir les figures 2.13 et 2.14, en utilisant une constellation 2+4+2 APSK, similaire à celle présente dans le standard DVB-S2(X). On mesure la variance conditionnelle des symboles reçus tout en faisant varier l'angle à l'origine de l'anneau médian (la QPSK).

Enfin, les différentes métriques permettant d'évaluer les performances sont présentées.

C.3 Etat de l'art des méthodes de prédistorsion

Dans ce chapitre, nous présentons les différentes méthodes de prédistorsion présentes dans l'état de l'art. Dans la littérature, les méthodes sont classiquement catégorisées par leur domaine de mise en œuvre soit au niveau des symboles soit au niveau du signal sur-échantillonné (après le filtre de mise en forme). Cette distinction fait sens car chacune de ces classes ont des propriétés propres. Par exemple, les méthodes "signal" peuvent compenser l'accroissement spectral après l'amplificateur, mais elles ont tendance à augmenter le support spectral sur le lien montant, *a contrario* des méthodes "symboles" qui ne peuvent compenser l'accroissement spectral mais n'augmentent pas le support spectral sur la voie montante.

Hors, certaines méthodes peuvent être appliquées de manière générique dans les deux domaines. C'est pourquoi nous avons décidé de catégoriser les méthodes selon le type de mise en œuvre. Nous avons donc les méthodes dites directes (ou non-itérative) et itérative. Les méthodes itératives correspondent aux méthodes où le traitement pour obtenir les données prédistordues sont mises à jour et traitées de manière itérée; tandis que pour les méthodes directes, les données prédistordues sont obtenues en une seule passe.

Dans la famille des méthodes directes, nous retrouvons:

Inversion à l'ordre P Méthode construisant la prédistorsion à partir d'un modélisation de Volterra de la chaîne, cette méthode est complexe et très dépendante du modèle sous-jacent.

Compensation globale Méthode basée sur la construction d'une table qui n'est pas adéquate avec les longues réponses impulsives ni les grandes constellations.

Méthodes adaptatives (ILA/DLA) Ces méthodes adaptent de manière itérative, une structure définie a priori (souvent une structure de Volterra), les performances vont dépendre de la structure choisi.

Réseaux neuronaux Les réseaux neuronaux présentent de bonnes performances et sont très intéressantes par leurs capacités de généralisation mais leurs performances dépendent de la structure choisie, notamment, les fonctions d'activation, le nombre de neurones et le nombres de couches.

Dans la famille des méthodes itératives, nous retrouvons:

"Small Variation Algorithm" met à jour une trame itérativement en modifiant un symbole à la fois à partir du modèle.

Méthodes basées sur le point fixe Ces méthodes ré-écrivent le problème de linéarisation sous forme d'un problème de point fixe. Si les hypothèses du théorème du point fixe sont vérifiées, alors le théorème nous garantit l'unicité de la solution et nous donne le moyen d'obtenir la solution. Sinon quelques itérations peuvent améliorer les performances.

Approximation successive L'approximation successive est une variation de la précédente méthode.

C.4 Prédistorsion itérative

Dans ce chapitre, nous présentons des nouvelles méthodes de prédistorsion itératives.

Dans un premier temps, nous proposons de dériver une méthode basée sur le point fixe ayant pour but de minimiser l'erreur quadratique moyenne. Afin de pouvoir se comparer avec l'état de l'art, nous proposons une version "signal" équivalente. Cette dernière permet d'identifier uniquement le transpondeur et, non plus, les filtres d'émission et de réception. Cette méthode est comparée à l'approximation successive, qui utilise plus d'hyper-paramètres pour des performances similaires.

Puis, dans un second mouvement, nous comparons la méthode proposée vis-à-vis de la version signal de la prédistorsion basée sur le théorème du point fixe que l'on trouve classiquement dans la littérature dans le cas où le canal est parfaitement connue et dans le cas où ce dernier est identifié.

Les performances de la version symbole sont moins bonnes en terme d'ACI vis-à-vis de la version signal. La tendance est inversée pour l'erreur quadratique moyenne. Globalement, l'écart entre les deux méthodes en terme de dégradation totale dans le cas où le canal est parfaitement connu est proche. Néanmoins, dans le cas identifié, les performances sont en faveur de la version symbole.

Le point permettant d'obtenir les meilleures performances en terme de dégradation totale, offre un compromis entre la puissance du signal reçue et la distorsion résiduelle. Ce point ne tombe donc pas dans les conditions du théorème, malgré cela, nous avons vu que la méthode permet d'améliorer les performances de linéarisation. Nous nous sommes donc intéressés à assurer les meilleures performances pour un nombre donné d'itérations ainsi, que de pallier au problème d'inadéquation du modèle du canal identifié avec le véritable canal.

Pour ce faire, nous avons introduit des hyper-paramètres de pondération et proposés une méthode d'adaptation.

L'introduction des hyper-paramètres permet d'obtenir des meilleurs performances en terme de NMSE et de dégradation totale; cependant cela introduit une faible perte en terme d'ACI.

Par la suite des perspectives d'extensions de la méthode sont évoquées, la diminution du facteur de crête et prédistorsion conjointe, une autre méthode d'adaptation, similaire à celle utilisé pour les réseaux neuronaux, et les risques de celle-ci, et enfin, une possibilité de solution en utilisant des paramètres d'adaptation variables.

C.5 Conclusion et perspectives

Le dernier chapitre synthétise les travaux effectués dans la thèse. Des perspectives sont évoquées, notamment, les liens potentiels avec les réseaux neuronaux, l'adaptation conjointe et une application possible pour le "Faster Than Nyquist".

Bibliography

- [1] A. C. Clarke, “Extra-terrestrial relays”, *Wirel. world*, vol. 51, pp. 305–308, 1945 (cit. on pp. [15](#), [16](#)).
- [2] European Space Agency, *EMERGSAT / ESA’s ARTES Programmes*, 2008 (cit. on p. [16](#)).
- [3] EUTELSAT, *Connect rural and remote GSM networks fast and cost-effectively* (cit. on p. [16](#)).
- [4] F. Khan, “Mobile Internet from the Heavens”, p. 8, 2015. arXiv: [1508.02383](#) (cit. on p. [16](#)).
- [5] A. Bauer, *Feu vert américain au projet de constellation de SpaceX*, 2018 (cit. on p. [16](#)).
- [6] European Space Agency, *Improving Internet access on high-speed trains / Space4Rail*, 2018 (cit. on p. [16](#)).
- [7] —, *Aeronautical Satellite Communications Channel Characteristics / ESA’s ARTES Programmes*, 2016 (cit. on p. [16](#)).
- [8] —, *A flexible new satcom-based marine communication system optimises bandwidth at sea / ESA’s ARTES Programmes*, 2014 (cit. on p. [16](#)).
- [9] C. Nicolas and C. Michel, *Satellite components for the 5G system*, 2018 (cit. on p. [16](#)).
- [10] European Telecommunications Standards Institute, “Digital Video Broadcasting (DVB): Framing structure, channel coding and modulation for 11/12 GHz satellite services”, *En 300 421*, vol. 1.1.2, pp. 1–24, 1997 (cit. on pp. [16](#), [27](#)).
- [11] —, *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 1: DVB-S2*, 2014 (cit. on pp. [16](#), [26](#), [29–31](#), [65](#)).
- [12] —, *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 2: DVB-S2 Extensions (DVB-S2X)*, 2015 (cit. on pp. [16](#), [26](#)).
- [13] G. Maral and M. Bousquet, *Satellite Communications Systems: Systems, Techniques and Technology*, 5th. Chichester, UK: John Wiley & Sons, Ltd, Dec. 2009 (cit. on pp. [16](#), [30](#), [33](#), [37](#)).
- [14] O. Gouba, “Joint Approach of Crest Factor Reduction and Linearization in OFDM context”, PhD thesis, Supélec, 2013 (cit. on p. [19](#)).

- [15] R. Piazza, M. Bhavani Shankar, and B. Ottersten, “Generalized direct predistortion with adaptive crest factor reduction control”, in *2015 IEEE Int. Conf. Acoust. Speech Signal Process.*, IEEE, Apr. 2015, pp. 3242–3246 (cit. on pp. 19, 52, 53, 69, 79).
- [16] T. Deleu, “Digital Predistortion and Equalization of the Non-Linear Satellite Communication Channel”, PhD thesis, Ecole Polytechnique, 2014 (cit. on p. 19).
- [17] S. Benedetto and E. Biglieri, “Nonlinear Equalization of Digital Satellite Channels”, *IEEE J. Sel. Areas Commun.*, vol. 1, no. 1, pp. 57–62, 1983 (cit. on p. 19).
- [18] M. Hotz and C. Vogel, “Linearization of time-varying nonlinear systems using a modified linear iterative method”, *IEEE Trans. Signal Process.*, vol. 62, no. 10, pp. 2566–2579, 2014. arXiv: [arXiv:1404.5901v1](https://arxiv.org/abs/1404.5901v1) (cit. on pp. 20, 56, 57, 63, 74, 75).
- [19] A. Ugolini, A. Modenini, G. Colavolpe, V. Mignone, and A. Morello, “Advanced techniques for spectrally efficient DVB-S2X systems”, *Int. J. Satell. Commun. Netw.*, vol. 34, no. 5, pp. 609–623, Sep. 2016 (cit. on p. 24).
- [20] A. Ginesi, “DVB-Sx Reference Simulation Scenarios for the Broadcast Profile”, 2013 (cit. on p. 25).
- [21] R. De Gaudenzi, A. Guillen i Fabregas, and A. Martinez, “Performance analysis of turbo-coded APSK modulations over nonlinear satellite channels”, *IEEE Trans. Wirel. Commun.*, vol. 5, no. 9, pp. 2396–2407, Sep. 2006 (cit. on pp. 26, 33, 37, 49).
- [22] S. Daumont, B. Rihawi, and Y. Louet, “Root-Raised Cosine filter influences on PAPR distribution of single carrier signals”, in *2008 3rd Int. Symp. Commun. Control Signal Process.*, IEEE, Mar. 2008, pp. 841–845 (cit. on p. 27).
- [23] H. Myung, “Single Carrier Orthogonal Multiple Access Technique for Broadband Wireless Communications”, PhD thesis, 2007 (cit. on p. 27).
- [24] D. Castelain, C. Ciochina, J. Guillet, and F. Hasegawa, “SC-OFDM, a Low-Complexity Technique for High Performance Satellite Communications”, *32nd AIAA Int. Commun. Satell. Syst. Conf.*, pp. 1–8, 2013 (cit. on p. 27).
- [25] D. Castelain and C. Ciochina-duchesne, “Combining Faster-than-Nyquist and SC-OFDM For increasing the spectral efficiency with a reasonable complexity”, (cit. on p. 27).
- [26] H. Ochiai, “Exact and Approximate Distributions of Instantaneous Power for Pulse-Shaped Single-Carrier Signals”, *IEEE Trans. Wirel. Commun.*, vol. 10, no. 2, pp. 682–692, Feb. 2011 (cit. on p. 28).
- [27] —, “On instantaneous power distributions of single-carrier FDMA signals”, *IEEE Wirel. Commun. Lett.*, vol. 1, no. 2, pp. 73–76, 2012 (cit. on p. 28).
- [28] A. A. M. Saleh, “Frequency-Independent and Frequency-Dependent Non-linear Models of TWT Amplifiers”, *IEEE Trans. Commun.*, vol. 29, no. 11, pp. 1715–1720, 1981 (cit. on pp. 31, 91).

- [29] D. Falconer, T. Kolze, Y. Leiba, and J. Liebetreu, "System Impairment Model", Tech. Rep., 2000 (cit. on pp. 31, 91).
- [30] B. F. Beidas, "Adaptive Digital Signal Predistortion for Nonlinear Communication Systems Using Successive Methods", *IEEE Trans. Commun.*, vol. 64, no. 5, pp. 2166–2175, May 2016 (cit. on pp. 31, 45, 57–60, 64–66).
- [31] S. Benedetto, E. Biglieri, and R. Daffara, "Modeling and Performance Evaluation of Nonlinear Satellite Links-A Volterra Series Approach", *Aerosp. Electron. Syst. IEEE Trans.*, vol. AES-15, no. 4, pp. 494–507, 1979 (cit. on pp. 33, 34).
- [32] R. Balasubramanian and M. Miller, "Pre-distortion techniques and bandwidth efficient modulation for military satellite communications", in *IEEE MILCOM 2004. Mil. Commun. Conf. 2004.*, vol. 2, IEEE, Oct. 2004, pp. 1028–1033 (cit. on pp. 33, 37).
- [33] D. Morgan, Z. Ma, J. Kim, M. G. Zierdt, and J. Pastalan, "A generalized memory polynomial model for digital predistortion of RF power amplifiers", *IEEE Trans. Signal Process.*, vol. 54, no. 10, pp. 3852–3860, 2006 (cit. on pp. 33, 50, 66).
- [34] F. Giri and E.-W. Bai, *Block-oriented Nonlinear System Identification*, F. Giri and E.-W. Bai, Eds., ser. Lecture Notes in Control and Information Sciences 1. London: Springer London, 2010, vol. 404, pp. 74–75 (cit. on p. 33).
- [35] G. Mzyk, *Combined Parametric- Nonparametric Identification of Block-Oriented Systems*. Springer, 2014 (cit. on p. 33).
- [36] V. Mathews, "Adaptive Polynomial Filters", *IEEE Signal Process. Mag.*, vol. 8, no. 3, pp. 10–26, 1991 (cit. on pp. 34, 35, 37).
- [37] M. Schetzen, *The Volterra and Wiener theories of nonlinear systems*. John Wiley & Sons,inc, 1980 (cit. on pp. 34, 48).
- [38] S. Benedetto and E. Biglieri, *Principles of Digital Transmission: With Wireless Applications*. Springer Science & Business Media, 1999 (cit. on p. 34).
- [39] J. Tsimbinos and K. Lever, "Computational complexity of Volterra based nonlinear compensators", *Electron. Lett.*, vol. 32, no. 9, p. 852, 1996 (cit. on p. 36).
- [40] L. Ding, "Digital Predistortion of Power Amplifiers for Wireless Applications", PhD thesis, Georgia Institute of Technology, 2004 (cit. on pp. 36, 50).
- [41] E. Casini, R. D. Gaudenzi, and A. Ginesi, "DVB-S2 modem algorithms design and performance over typical satellite channels", *Int. J. Satell. Commun. Netw.*, vol. 22, no. 3, pp. 281–318, 2004 (cit. on pp. 37, 48, 49, 62).
- [42] B. Picinbono and P. Chevalier, "Widely linear estimation with complex data", *IEEE Trans. Signal Process.*, vol. 43, no. 8, pp. 2030–2033, 1995 (cit. on p. 41).
- [43] M. Schetzen, "Theory of pth-order inverses of nonlinear systems", *IEEE Trans. Circuits Syst.*, vol. 23, no. 5, pp. 285–291, May 1976 (cit. on p. 48).

- [44] A. Sarti and S. Pupolin, "Recursive techniques for the synthesis of a pth-order inverse of a volterra system", *Eur. Trans. Telecommun.*, vol. 3, no. 4, pp. 315–322, 1992 (cit. on p. 49).
- [45] G. Karam and H. Sari, "A data predistortion technique with memory for QAM radio systems", *IEEE Trans. Commun.*, vol. 39, no. 2, pp. 336–344, 1991 (cit. on p. 49).
- [46] Changsoo Eun and E. Powers, "A new Volterra predistorter based on the indirect learning architecture", *IEEE Trans. Signal Process.*, vol. 45, no. 1, pp. 223–227, 1997 (cit. on pp. 49, 50).
- [47] A. A. M. Saleh and J. Salz, "Adaptive Linearization of Power Amplifiers in Digital Radio Systems", *Bell Syst. Tech. J.*, vol. 62, no. 4, pp. 1019–1033, Apr. 1983 (cit. on p. 49).
- [48] L. Ding, R. Raich, and G. Zhou, "A Hammerstein predistortion linearization design based on the indirect learning architecture", *IEEE Int. Conf. Acoust. Speech Signal Process.*, vol. 3, no. 4, pp. III–2689–III–2692, 2002 (cit. on p. 50).
- [49] R. Dallinger, H. Ruotsalainen, R. Wichman, and M. Rupp, "Adaptive predistortion techniques based on orthogonal polynomials", *2010 Conf. Rec. Forty Fourth Asilomar Conf. Signals, Syst. Comput.*, pp. 1945–1950, 2010 (cit. on p. 50).
- [50] E. Abd-Elrady and Li Gan, "Adaptive predistortion of Hammerstein systems based on indirect learning architecture and prediction error method", in *2008 Int. Conf. Signals Electron. Syst.*, IEEE, Sep. 2008, pp. 389–392 (cit. on pp. 50, 54).
- [51] Anding Zhu and T. Brazil, "Adaptive Volterra-based predistorter design for RF high power amplifiers", in *6th IEEE High Freq. Postgrad. Colloq. (Cat. No.01TH8574)*, IEEE, 2001, pp. 100–105 (cit. on p. 50).
- [52] R. Raich, H. Qian, and G. Zhou, "Orthogonal polynomials for power amplifier modeling and predistorter design", *IEEE Trans. Veh. Technol.*, vol. 53, no. 5, pp. 1468–1479, 2004 (cit. on p. 50).
- [53] H. W. Kang, Y. S. Cho, and D. H. Youn, "Adaptive precompensation of wiener systems", *IEEE Trans. Signal Process.*, vol. 46, no. 10, pp. 2825–2829, 1998 (cit. on p. 50).
- [54] D. Zhou and V. E. DeBrunner, "Novel adaptive nonlinear predistorters based on the direct learning algorithm", *IEEE Trans. Signal Process.*, vol. 55, no. 1, pp. 120–133, 2007 (cit. on pp. 50, 51, 69).
- [55] B. Widrow and E. Walach, *Adaptive Inverse Control, Reissue Edition: A Signal Processing Approach*. Wiley-IEEE Press, 2008 (cit. on pp. 50, 52).
- [56] B. Widrow and S. D. Stearns, *Adaptive signal processing*. 1985, vol. 1 (cit. on p. 50).
- [57] S. J. Elliott, "Filtered reference and filtered error LMS algorithms for adaptive feedforward control", *Mech. Syst. Signal Process.*, vol. 12, no. 6, pp. 769–781, 1998 (cit. on pp. 51, 52).
- [58] R. Zayani and R. Bouallegue, "Pre-Distortion for the compensation of HPA nonlinearity with neural networks : Application to satellite communications", vol. 7, no. 3, pp. 97–103, 2007 (cit. on pp. 53, 54).

- [59] R. Zayani, R. Bouallegue, and D. Roviras, “Adaptive predistortions based on neural networks associated with Levenberg-Marquardt algorithm for satellite down links”, *Eurasip J. Wirel. Commun. Netw.*, vol. 2008, 2008 (cit. on pp. 53, 54).
- [60] T. Gotthans, G. Baudoin, and A. Mbaye, “Digital predistortion with advance/delay neural network and comparison with Volterra derived models”, *2014 IEEE 25th Annu. Int. Symp. Pers. Indoor, Mob. Radio Commun.*, pp. 811–815, 2014 (cit. on pp. 53, 54, 79).
- [61] M. Ibnkahla, “Neural network predistortion technique for digital satellite communications”, *ICASSP, IEEE Int. Conf. Acoust. Speech Signal Process. - Proc.*, vol. 6, no. section 3, pp. 3506–3509, 2000 (cit. on pp. 53, 54).
- [62] H. Abdulkader, F. Langlet, D. Roviras, and F. Castanie, “Natural gradient algorithm for neural networks applied to non-linear high power amplifiers”, *Int. J. Adapt. Control Signal Process.*, vol. 16, no. 8 SPEC. Pp. 557–576, 2002 (cit. on pp. 53, 54).
- [63] M. Rawat, K. Rawat, and F. M. Ghannouchi, “Adaptive digital predistortion of wireless power amplifiers/transmitters using dynamic real-valued focused time-delay line neural networks”, *IEEE Trans. Microw. Theory Tech.*, vol. 58, no. 1, pp. 95–104, 2010 (cit. on pp. 53, 54).
- [64] S. Haykin, *Neural Networks: A Comprehensive Foundation*. Prentice Hall, 1999 (cit. on pp. 53, 54, 80).
- [65] G. Cybenko, “Approximation by superpositions of a sigmoidal function”, *Math. Control. Signals, Syst.*, vol. 2, no. 4, pp. 303–314, Dec. 1989 (cit. on p. 54).
- [66] F. Mkadem and S. Boumaiza, “Physically inspired neural network model for RF power amplifier behavioral modeling and digital predistortion”, *IEEE Trans. Microw. Theory Tech.*, vol. 59, no. 4 PART 1, pp. 913–923, 2011 (cit. on p. 54).
- [67] N. Benvenuto, M. Marchesi, F. Piazza, and A. Uncini, *A comparison between real and complex-valued neural networks in communication applications*, 1991 (cit. on p. 54).
- [68] W. Liu, J. C. Principe, and S. Haykin, *Kernel Adaptive Filtering*. Hoboken, NJ, USA: John Wiley & Sons, Inc., Feb. 2010 (cit. on p. 54).
- [69] T. Deleu, M. Dervin, K. Kasai, and F. Horlin, “Iterative Predistortion of the nonlinear satellite channel”, *IEEE Trans. Commun.*, vol. 62, no. 8, pp. 2916–2926, Aug. 2014. arXiv: [arXiv:1401.4944v1](https://arxiv.org/abs/1401.4944v1) (cit. on pp. 55, 56, 64).
- [70] R. Nowak and B. Van Veen, “Volterra filter equalization: a fixed point approach”, *IEEE Trans. Signal Process.*, vol. 45, no. 2, pp. 377–388, Feb. 1997 (cit. on pp. 56, 57, 63, 74, 75).
- [71] E. Aschbacher, M. Steinmair, and M. Rupp, “Iterative linearization methods suited for digital pre-distortion of power amplifiers”, *Conf. Rec. Thirty-Eighth Asilomar Conf. Signals, Syst. Comput. 2004.*, vol. 2, no. 4, pp. 5–9, 2004 (cit. on pp. 56, 57, 74, 75).

- [72] M.-C. Kim, Y. Shin, and S. Im, “Compensation of nonlinear distortion using a predistorter based on the fixed point approach in OFDM systems”, *VTC '98. 48th IEEE Veh. Technol. Conf.*, vol. 3, pp. 2145–2149, 1998 (cit. on pp. [57](#), [63](#), [74](#), [75](#)).
- [73] E. Zeidler, *Nonlinear functional analysis and its application I: Fixed-Point Theorems*. Springer-Verlag New York, 1985 (cit. on p. [57](#)).
- [74] J. E. Mazo, “Faster-than-Nyquist signaling”, *Bell Syst. Tech. Journal*, vol. 54, no. 8, pp. 1451–1462, 1975 (cit. on pp. [63](#), [85](#)).
- [75] J.-a. Lucciardi, N. Thomas, M.-l. Boucheret, and C. Poulliat, “Trade-Off Between Spectral Efficiency Increase And PAPR Reduction When Using FTN Signaling : Impact Of Non Linearities”, 2016 (cit. on p. [63](#)).
- [76] L. Ding, G. Zhou, D. Morgan, Z. Ma, J. S. Kenney, J. Kim, and C. R. Giardina, “A Robust Digital Baseband Predistorter Constructed Using Memory Polynomials”, *IEEE Trans. Commun.*, vol. 52, no. 1, pp. 159–165, 2004 (cit. on pp. [66](#), [67](#)).
- [77] T. Ogunfunmi, *Adaptive Nonlinear System Identification*, ser. Signals And Communication Technology. Boston, MA: Springer US, 2007 (cit. on p. [66](#)).
- [78] B. Huang, Y. Xiao, J. Sun, and G. Wei, “A variable step-size FXLMS algorithm for narrowband active noise control”, *IEEE Trans. Audio, Speech Lang. Process.*, vol. 21, no. 2, pp. 301–312, 2013 (cit. on p. [80](#)).
- [79] M. Raghu, J. Kleinberg, and B. Poole, “On the expressive power of deep neural networks”, arXiv: [1611.08083v1](#) (cit. on p. [84](#)).
- [80] F. Rusek and J. Anderson, “The two dimensional Mazo limit”, in *Proceedings. Int. Symp. Inf. Theory, 2005. ISIT 2005.*, vol. 0, IEEE, 2005, pp. 970–974 (cit. on p. [85](#)).
- [81] P. J. Schreier and L. L. Scharf, *Statistical signal processing of complex-valued data: The theory of improper and noncircular signals*. 2010, vol. 9780521897, pp. 1–309. arXiv: [arXiv:1011.1669v3](#) (cit. on p. [88](#)).

Titre : Algorithmes itératifs de prédistorsion adaptés à la montée en débit des communications par satellite.

Mots clés : Prédistorsion, Satellite, Amplificateur

Résumé : La thèse concerne les techniques de prédistorsion appliquées aux communications par satellite de type DVB-S2X dans la perspective de la montée en débit de ces systèmes. Pour atteindre cet objectif, de faibles facteurs de retombée (inférieurs à 5%) et des ordres de modulation plus élevés sont préconisés, augmentant la vulnérabilité vis-à-vis des non-linéarités introduites lors du passage par le satellite. Dans la littérature, deux classes de méthodes se distinguent par leur performance de linéarisation. La première est constituée des architectures d'apprentissage direct (Direct Learning Architecture) et indirect (Indirect Learning Architecture) et la seconde regroupe les méthodes itératives basées sur le théorème du point fixe.

Le principal défaut de la première classe est la nécessité de choisir une structure de prédistorsion a priori tandis que pour la deuxième classe, les conditions d'application du théorème du point fixe sont difficiles à vérifier dans la majorité des cas et nécessitent l'addition d'un gain empirique.

Nous avons donc proposé 1) un schéma de prédistorsion itérative basé sur le théorème du point fixe, appliqué sur le signal mais optimisé en minimisant une erreur quadratique entre les symboles de modulation et la sortie de l'échantillonneur, 2) un schéma de prédistorsion itérative incluant une adaptation automatique des paramètres et adapté à tout type de canal.

Title : Iterative predistortion algorithms adapted to the increasing throughput of satellite communications.

Keywords : Predistortion, Satellite, Amplifier

Abstract : The thesis focuses on predistortion techniques applied to DVB-S2X satellite communications in order to increasing the throughput of these systems. To achieve this objective, sharp roll-off factors (less than 5%) and higher modulation orders are recommended, increasing sensibility with regard to the non-linearities introduced during the passage by the satellite. In the literature, two classes of methods are distinguished by their linearization performance. The first consists of direct (Direct Learning Architecture) and indirect (Indirect Learning Architecture) learning architectures and the latter includes iterative methods based on the fixed point theorem.

The main defect of the first class is the need to choose a predistortion structure a priori, while for the second class, the conditions for applying the fixed point theorem are difficult to verify in most cases and require the addition of an empirical gain.

We therefore proposed 1) an iterative predistortion scheme based on the fixed point theorem, applied to the signal but optimized by minimizing a quadratic error between the modulation symbols and the sampler output, 2) an iterative predistortion scheme including an automatic adaptation of the parameters and suitable for all types of channels.