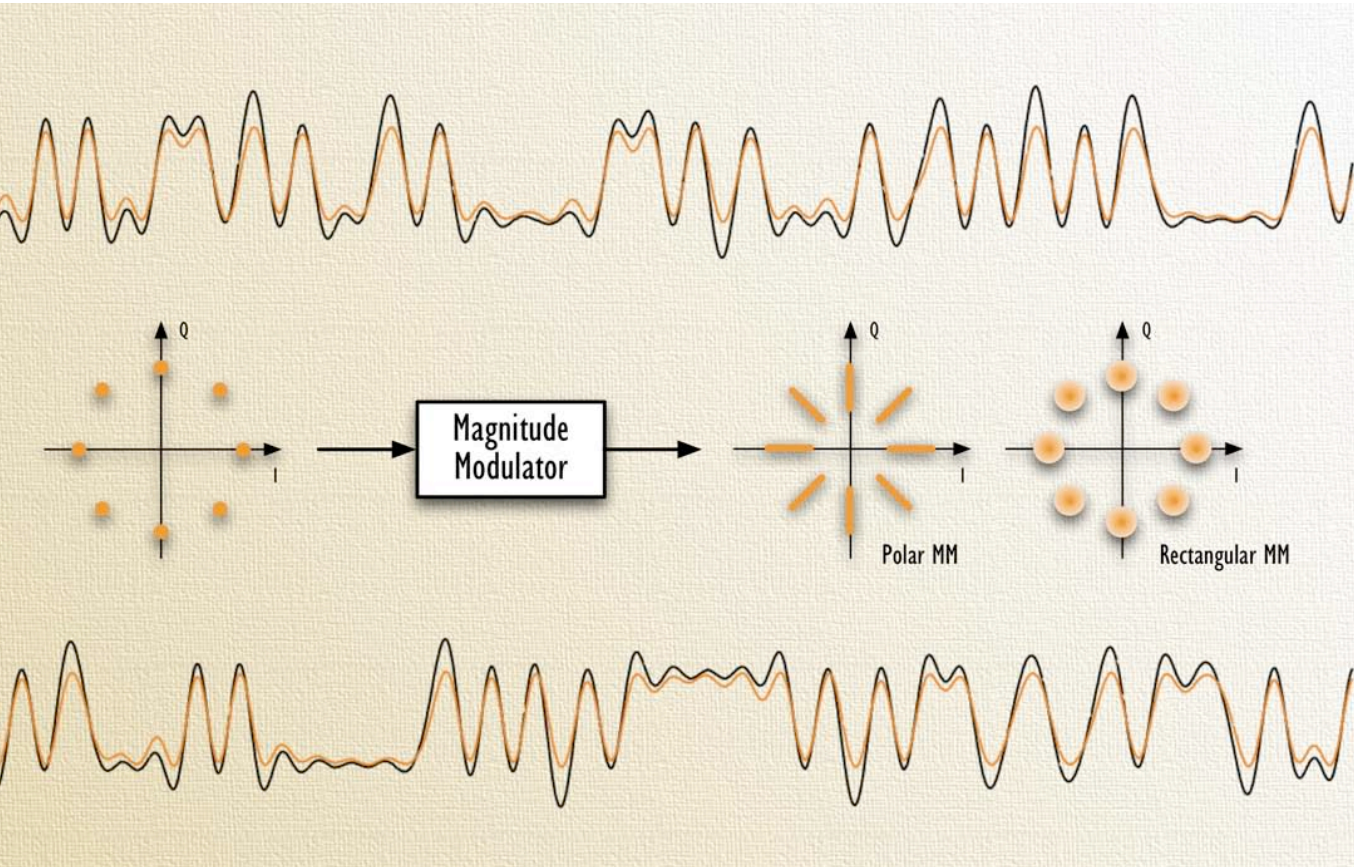


Marco Alexandre Cravo Gomes

Magnitude Modulation for Peak Power Control in Single Carrier Communication Systems

Dissertação de Doutoramento na área científica de Engenharia Electrotécnica e de Computadores, especialidade de Telecomunicações e Electrónica, orientada pelo Senhor Professor Doutor Vitor Manuel Mendes da Silva e pelo Senhor Professor Doutor Francisco António Bucho Cercas e apresentada ao Departamento de Engenharia Electrotécnica e de Computadores da Faculdade de Ciências e Tecnologia da Universidade de Coimbra.

Dezembro de 2010

· U C ·

Universidade de Coimbra

FCTUC FACULDADE DE CIÊNCIAS
E TECNOLOGIA
UNIVERSIDADE DE COIMBRA

MAGNITUDE MODULATION FOR PEAK POWER CONTROL IN SINGLE CARRIER COMMUNICATION SYSTEMS

MARCO ALEXANDRE CRAVO GOMES

Thesis submitted in partial fulfillment of the requirements for the Degree of Doctor of Philosophy in Electrical and Computer Engineering, specialization of Telecommunications and Electronics

Tese submetida para a satisfação dos requisitos para obtenção do grau de Doutor em Engenharia Electrotécnica e de Computadores, especialização em Telecomunicações e Electrónica

Coimbra
December 2010

THIS THESIS WORK HAS BEEN PARTIALLY SUPPORTED BY FUNDAÇÃO PARA A CIÊNCIA E A TECNOLOGIA (FCT), UNDER PHD GRANT **SFRH/BD/**38338/2007.

SUPERVISION

VITOR MANUEL MENDES DA SILVA

Professor Auxiliar do Departamento de Engenharia Electrotécnica e de Computadores da FCTUC

FRANCISCO ANTÓNIO BUCHO CERCAS

Professor Associado com Agregação no Departamento de Ciências e Tecnologias da Informação do ISCTE-IUL

Amo como ama o amor.
Não conheço nenhuma outra razão para amar senão amar.
Que queres que te diga, além de que te amo,
se o que quero dizer-te é que te amo?
— Fernando Pessoa

À Sara e aos meus Pais...
Dedicated to Sara and my Parents.

I don't know what tomorrow will bring...

— Fernando Pessoa

ACKNOWLEDGMENTS

A frase de Fernando Pessoa define bem a nossa essência, a incerteza do amanhã... É assim também em Ciência. Penso ser esse o espírito de quem inicia o seu trabalho de doutoramento, tal como eu, em 2006. Move-nos, no entanto, a nossa busca incessante pelo Saber, o descobrir do amanhã.

Pelo caminho, há pessoas que directa ou indirectamente nos vão ajudando a definir o rumo, a alcançar pequenas vitórias (portos de abrigo) antes de enfrentar de novo a incerteza. É a essas pessoas, e também às instituições que me forneceram os meios para percorrer este caminho, que aqui agradeço...

Agradeço ao Instituto de Telecomunicações, por todos os meios materiais e financeiros disponibilizados, e à Fundação para a Ciência e Tecnologia, por ter financiado em parte este trabalho de doutoramento.

Aos meus amigos e às pessoas do DEEC e do IT (professores, alunos, funcionários e investigadores), por todo o seu apoio. Um agradecimento especial ao colega Gabriel Falcão que comigo colaborou activamente ao longo dos últimos anos, pelo seu espírito empreendedor e entusiasta. Perdoem-me todos os outros, por não referir o vosso nome, mas a lista seria interminável...

Special thanks are due to Professor Martin Tomlinson for his wise knowledge and sympathy, from whom I learned a lot during our discussions, and to his research group, for receiving me so well, making me feel at home, and helping me to cope with the loneliness of being away from my wife and my family. I left good friends in Plymouth...

Ao Professor Francisco Cercas pela sua sábia orientação e disponibilidade, e pelas sempre confortantes e constantes palavras de incentivo.

Ao Professor Vitor Silva. Mais do que um orientador, um Amigo, "como da família" (tal como ele me disse um dia). Palavras que nos marcam e que para sempre irei guardar. Desde cedo um tutor, com a sua infinita dedicação...

À minha família, o meu refúgio, o meu ser...

À Sara pelo seu amor, carinho, ternura, amizade, cumplicidade, infinita compreensão, incentivo... por tudo!

A todos,

Muito Obrigado.

ABSTRACT

Peak-to-average power ratio (PAPR) reduction is of major concern in most telecommunication systems whose transmission capacity is limited by the transmitter's non-linear *high-power amplifier* (HPA). A high PAPR puts stringent linearity requirements on the HPA, which badly degrades the system's power efficiency. In addition it also complicates the design of other components of the system, e.g. the *digital-to-analog converter* (DAC), thereby increasing the overall cost of the system.

This thesis discusses the problem of controlling the envelope's power peak of *single-carrier* (SC) modulated signals, band limited by *Nyquist pulse shaping filters*, usually achieved with *root-raised cosine* (RRC) filters. Our main aim is to lower the signal's PAPR in order to reduce the power amplifier *back-off* for *small satellite communication terminals* (VSAT).

Following original work by Miller *et al.* and Tomlinson *et al.* on the PAPR controlling of SC signals we have focused on the concept of magnitude modulation (MM). The MM technique is presented as a very efficient solution to the PAPR problem. This thesis contributes two highly efficient MM algorithms that control the envelope's power peak: an improved LUT-based approach where constellation and RRC symmetries are explored, and a novel multistage polyphase magnitude modulation (MPMM) approach. The increased sensitivity to noise due to MM is also addressed, with two effective techniques being proposed for improving the detection of MM signals. Another major contribution is an analytical model for statistically characterizing and analyzing the performance of the MM procedure.

SUMÁRIO

A redução da relação entre potência de pico e a potência média (PAPR) de um sinal é motivo de preocupação no projecto da grande maioria de sistemas de telecomunicações, cuja capacidade de transmissão é limitada pela não linearidade do amplificador de potência (HPA) usado no emissor. Um elevado PAPR coloca elevados requisitos de linearidade no HPA, com consequente perda de eficiência de potência. Para além disso, dificulta igualmente o projecto de outros componentes do sistema, tais como o conversor digital-analógico (DAC), contribuindo para o aumento de custo global do sistema.

Esta tese aborda o problema de controlo da potência de pico da envolvente de sinais modulados em portadora única e com largura de banda espectral limitada com recurso a um filtro de Nyquist (tipicamente um filtro RRC). O nosso principal objectivo é minimizar o PAPR do sinal a transmitir, com vista a reduzir a necessidade de *back-off* em sistemas VSAT.

Dando continuidade ao trabalho de Miller *et al.* e Tomlinson *et al.*, concentramo-nos no estudo do conceito de modulação de magnitude (MM). A técnica de MM é apresentada como uma solução muito eficiente para a resolução do problema associado a um elevado PAPR. Este trabalho contribui com dois algorítmos de MM com elevado desempenho no controlo da potência de pico da envolvente do sinal a transmitir, sendo proposta uma versão melhorada do método de MM baseada em LUT's, bem como uma nova técnica polifásica de modulação de magnitude (MPMM). O problema de aumento de sensibilidade ao ruído, devido ao uso de técnicas de MM, é também abordado. Assim, este trabalho propõe duas novas técnicas que melhoram eficazmente a detecção do sinal modulado em magnitude. Outra contribuição muito importante é o desenvolvimento de um modelo analítico para caracterização e análise estatística do desempenho do método de MM.

PUBLICATIONS

The original contributions presented in this thesis were disseminated in the following scientific publications.

SCIENTIFIC JOURNALS WITH REFEREES:

M. Gomes, F. Cercas, V. Silva, and M. Tomlinson, "Magnitude modulation for VSAT's low back-off transmission," *Journal Commun. Networks (JCN), special issue on Recent Adv. in Satell. and Space Commun.*, vol. 12, no. 6, pp. 544–557, Dec. 2010.

M. Gomes, V. Silva, F. Cercas, and M. Tomlinson, "Power efficient back-off reduction through polyphase filtering magnitude modulation," *IEEE Commun. Lett.*, vol. 13, no. 8, pp. 606–608, August 2009.

SCIENTIFIC CONFERENCES WITH REFEREES:

M. Gomes, V. Silva, F. Cercas, and M. Tomlinson, "Analytical analysis of polyphase magnitude modulation method's performance," in *Proc. IEEE ICC'2010*, Cape Town, South Africa, May 2010.

M. Gomes, F. Cercas, V. Silva, and M. Tomlinson, "Polyphase magnitude modulation for peak power control," in *17th European Signal Processing Conference, EUSIPCO 2009.*, Glasgow, Scotland, August 2009.

M. Gomes, V. Silva, F. Cercas, and M. Tomlinson, "Real-time LUT-less magnitude modulation for peak power control of single carrier rrc filtered signals," in *IEEE 10th Workshop on Signal Process. Advances in Wireless Commun., SPAWC '09.* , Perugia, Italy, June 2009, pp. 424–428.

M. Gomes, F. Cercas, V. Silva, and M. Tomlinson, "Efficient m-QAM transmission using compacted magnitude modulation tables," in *Proc. IEEE Globecom'08*, New Orleans, LA, Nov. 2008.

M. Gomes, V. Silva, F. Cercas, and M. Tomlinson, "Low back-off 16-APSK transmission using magnitude modulation and symbol quantization," in *IEEE Int. Workshop on Satellite and Space Commun., IWSSC'08.* , Toulouse, France, Oct. 2008, pp. 229–233.

M. Gomes, G. Falcao, V. Silva, V. Ferreira, A. Sengo, L. Silva, N. Marques, and M. Falcao, "Scalable and parallel codec architectures for the dvb-s2 fec system," in *Proc. IEEE APCCAS'08*, Macau, China, Nov. 2008, pp. 1506–1509.

M. Gomes, G. Falcao, A. Sengo, V. Ferreira, V. Silva, and M. Falcao, "High throughput encoder architecture for DVB-S2 LDPC-IRA codes," in *Proc. IEEE Int. Conf. Microelectronics, ICM'07*, Cairo, Egypt, Dec. 2007, pp. 271–274.

M. Gomes, G. Falcao, V. Silva, V. Ferreira, A. Sengo, and M. Falcao, "Flexible parallel architecture for DVB-S2 LDPC decoders," in *Proc. IEEE Globecom'07*, Washington, DC, Nov. 2007, pp. 3265–3269.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACRONYMS

ADC	Analog-to-Digital Converter
AGC	Automatic Gain Control
APP	A Posteriori Probability
APSK	Amplitude Phase-Shift Keying
ARQ	Automatic Repeat Request
AWGN	Additive White Gaussian Noise
BCH	Bose-Chaudhuri-Hocquenghem Code
BER	Bit Error Rate
BLER	Block Error Rate
BN	Bit Node
BO	Back-off
BPSK	Binary Phase-Shift Keying
BSC	Binary Symmetric Channel
CCDF	Complementary Cumulative Distribution Function
CD	Compact Disc
CN	Check Node
CRC	Cyclic Redundancy Check Code
DAC	Digital-to-Analog Converter
D/A	Digital-to-Analog
DTFT	Discrete-Time Fourier Transform
DVB-S2	Digital Video Broadcast - Satellite (*second generation*)
EAC	Expected Average Constellation
EVM	Error Vector Magnitude
FEC	Forward Error Correction

FIR	Finite Impulse Response
FU	Functional Unit
GF	Galois Field
GSM	Global System for Mobile Communications
HPA	High Power Amplifier
IBO	Input Back-off
i.i.d.	independent and identically distributed
IN	Information Nodes
IRA	Irregular Repeat-Accumulate Code
ISI	Intersymbol Interference
LDPC	Low-Density Parity-Check Code
LLR	Log-Likelihood Ratio
LTE	Long Term Evolution
LTI	Linear Time-Invariant
LUT	Look-Up Table
LUT-MM	LUT-based MM implementation
MAP	Maximum a Posteriori Probability
MBWA	Mobile Broadband Wireless Access
MC	Multi-Carrier
ML	Maximum Likelihood
MM	Magnitude Modulation
MODCOD	Modulation-Codification Scheme
MPMM	Multistage Polyphase Magnitude Modulation
MPMM+SD	Join MPMM *plus* Soft-Decoding Detection
OBO	Output Back-off
OFDM	Orthogonal Frequency-Division Multiplexing
OQPSK	Offset QPSK

PAPR	Peak-to-Average Power Ratio
PC-RS	LUT-MM Polar-Clipping Rectangular-Scaling Case
pdf	*probability density function*
PN	Parity Nodes
PSK	Phase-Shift Keying
QAM	Quadrature Amplitude Modulation
QPSK	Quadrature Phase-Shift Keying
RAM	Random Access Memory
r.v.	random variable
RC	Raised Cosine
RC-RS	LUT-MM Rectangular-Clipping Rectangular-Scaling Case
ROM	Read Only Memory
RS	Reed-Solomon Code
RRC	Root-Raised Cosine
SC	Single-Carrier
SC-FDMA	Single-Carrier Frequency Division Multiple Access
SER	Symbol Error Rate
SNR	Signal-to-Noise Ratio
SPA	Sum-Product Algorithm
SSS	Strick Sense Stationary
SWR	Software Radio
TC	Trellis-Coding
Tx	Transmitter
VSAT	Very-Small Aperture Terminals
WSCS	Wide Sense Cyclostationary
WSS	Wide Sense Stationary
WiMAX	Worldwide Interoperability for Microwave Access

1 INTRODUCTION

1.1 MOTIVATION OF THE THESIS

This thesis discusses the problem of controlling the envelope's power peak of *single-carrier* (SC) modulated signals, band-limited by *Nyquist pulse shaping filters* [44, 65, 87], usually achieved with *root-raised cosine* (RRC) filters. The main aim is to reduce power amplifier *back-off* for *small satellite communication terminals* (VSAT) [53, 70].

Peak-to-average power ratio (PAPR) reduction has always been a concern in bandwidth limited channels, e.g. satellite channels, whose transmission capacity is limited by the transmitter's non-linear *high-power amplifier* (HPA). A high PAPR puts stringent requirements on the transmitter's *digital-to-analog converter* (DAC) and on the HPA, increasing the cost of these systems' components and badly degrading the system's power efficiency.

In recent years much attention has been paid to solving the problem of reducing the PAPR in multi-carrier modulation systems, such as *orthogonal frequency-division multiplexing* (OFDM) [42, 57, 67]. But achieving low PAPR with strictly band-limited SC signals is also desirable, especially with the renewed interest in SC as the preferred solution to be used on the *uplink* by the next-generation wireless communications standards, e.g. LTE [46, 95]. The ability to combine the lower PAPR of traditional SC formats with the multi-path resistance and in-channel frequency scheduling flexibility of OFDM in hybrid modulation schemes such as *single-carrier frequency division multiple access* (SC-FDMA) [95], led to the decision to choose SC for the uplink transmission in order to save the precious battery life of small mobile devices.

However, the apparently *"simpler"* problem of achieving low PAPR with strictly band-limited SC signals, is still unsolved. Although the peak power problem seems to be less critical in SC modulations than in OFDM, it is also true that the SC PAPR can be extremely high if a very low *roll-off* Nyquist filter is used [96]. This problem has particular relevance in VSAT systems. With the aim of minimizing terminals' power consumption, and so the need to reduce the transmitted signal's PAPR, constant amplitude SC modulation formats, such as *phase-shift keying* (PSK), and pulse shaping with moderate roll-off (typically $\alpha = 0.35$) have always been the preferred choice for VSAT [54, 70]. But in order to meet the growing demands for higher data rates (and so higher *spectral efficiency*), higher-order non-constant amplitude constellations and sharper Nyquist filtering are now being used. As an example, in the

recent satellite standard for *Digital Video Broadcasting* (DVB-S2) [19], new M-APSK constellations were added [15] (with $M = \{16, 32\}$) supported by a set of spectrum shapes with roll-off factors of $\alpha = \{0.2, 0.25, 0.35\}$. These lead to a considerable increase in the PAPR of the transmitted signal, emphasizing the peak power problem.

The contributions to the PAPR of a band-limited SC signal, and the effect of the PAPR on the performance of an HPA, are discussed in more detail below.

1.2 THE PEAK POWER PROBLEM

The generic block diagram of a typical single carrier (SC) transmitter is shown in Fig. 1. The following processing stages are employed to obtain the transmit signal $t_x(t)$. Redundancy defined by an *error-correcting code* is appended to the information by the channel encoder so that most errors occurring during transmission can be corrected [66, 77]. The encoded data is then mapped into complex symbols $s[n]$ of the constellation for the modulation scheme in use [10, 87]. According to Nyquist theory [44, 76], the modulated signal has to be up-sampled by $L \geqslant 2$, before limiting its bandwidth. A Nyquist pulse shaping filter limits the bandwidth of the up-sampled modulated signal $s_u[n]$ in order to guarantee transmission over the band-limited channel without *intersymbol interference* (ISI) [44]. The shaped sequence is converted to a continuous-time signal in the DAC [44, 76] and amplified by the HPA [59, 67]. Finally, if necessary, the baseband signal becomes passband by translation to a higher frequency [10, 87].

In SC systems, the main contribution to PAPR results from RRC bandwidth limitation[1] that gives rise to undesirable envelope variations in the transmitted signal $x[n]$ that drives the transmitter's DAC+HPA block.

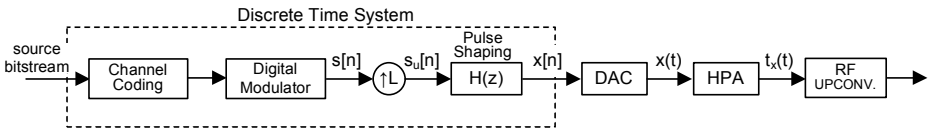


Figure 1: Generic SC transmitter's system block diagram.

The equivalent complex digital baseband signal, after filtering, and prior to DAC conversion, is given by:

$$x[n] = \left[\sum_k s[k]\, \delta[n - kL] \right] * h[n] \ , \tag{1.1}$$

1 Hereafter, we will consider, without loss of generalization, that Nyquist pulse shaping is achieved with RRC filters, as it usually is in practice. Hence, both terms will be used interchangeably throughout the text, although the term *Nyquist filter* refers to a broader class of filters to which the techniques that will be discussed throughout this thesis also apply.

where $h[n]$ is the impulse response of the pulse shaping RRC filter and $L \in \mathbb{N}$ is the oversampling factor.

The PAPR of the transmitted signal $x[n]$ is defined as

$$\text{PAPR} = 10\log_{10}\left[\frac{\max|x[n]|^2}{E\left\{|x[n]|^2\right\}}\right] \quad (\text{dB}) , \tag{1.2}$$

and it is the sum of two components (as shown in appendix A): $\text{PAPR}_{\text{Const}}$ and PAPR_{RRC}, due to the constellation and the RRC filter, respectively, which are evaluated as

$$\text{PAPR}_{\text{Const}} = 10\log_{10}\left[\frac{\max\limits_{i=1,\cdots,M}|S_i|^2}{\frac{1}{M}\sum\limits_{i=1}^{M}|S_i|^2}\right] \quad (\text{dB}) , \tag{1.3}$$

$$\text{PAPR}_{\text{RRC}} = 10\log_{10}\left[\frac{\max\limits_{i=0,\cdots,L-1}\left(\sum\limits_{n=0}^{2N}|h[nL+i]|\right)^2}{\frac{1}{L}\sum\limits_{n=0}^{2NL}|h[n]|^2}\right] \quad (\text{dB}) , \tag{1.4}$$

where S_i, with $i = 1, \cdots, M$ are the symbols of a M-ary constellation, and it is assumed, without loss of generalization, a type I linear phase FIR RRC filter [76] with $2NL + 1$ coefficients.

$\text{PAPR}_{\text{Const}}$ depends only on the constellation's geometry. It is clear from (1.3) that $\text{PAPR}_{\text{Const}}$ is null for constant amplitude constellations, such as M-PSK. However, for higher order modulations, it contributes with a non-negligible value to the total PAPR value, as shown in table 1.

Nevertheless, the major contribution to PAPR comes from Nyquist pulse shaping, especially when considering roll-off (α) values in the range of interest, $\alpha \leqslant 0.3$. 2 typical values are presented in table 2. The main goal of any PAPR reduction technique for SC transmission is, therefore, to cancel this undesirable PAPR_{RRC} contribution, eliminating the need to back-off from HPA saturation.

1.2.1 *Power amplifier*

A high PAPR puts stringent requirements on the linearity of the transmitter's HPA. But because of cost, design, and, most importantly, power efficiency constraints, the HPA is not perfectly linear. This causes in-band distortion and, even worse, *spectral spreading*, i.e. adjacent channel interference.

Constellation	M-PSK	16-APSK[a]	16-QAM	32-APSK[b]	64-QAM
$PAPR_{Const}$	0 dB	1.1 dB	2.6 dB	2.1 dB	3.7 dB

a. 16-APSK DVB-S2 constellation with $\gamma = 3.15$ [19].
b. 32-APSK DVB-S2 constellation with $\gamma_1 = 2.84$ and $\gamma_2 = 5.27$ [19].

Table 1: PAPR contribution from constellation.

Roll-off	0.1	0.2	0.3	0.4	0.5
$PAPR_{RRC}$	7.5 dB	5.8 dB	4.6 dB	3.7 dB	3.4 dB

Table 2: PAPR contribution from the RRC digital filter.

The nonlinear behavior of an HPA is usually characterized by the amplitude modulation/amplitude modulation (AM/AM) and amplitude modulation/phase modulation (AM/PM) conversion. Figure 2 shows a typical AM/AM characteristic of an HPA.

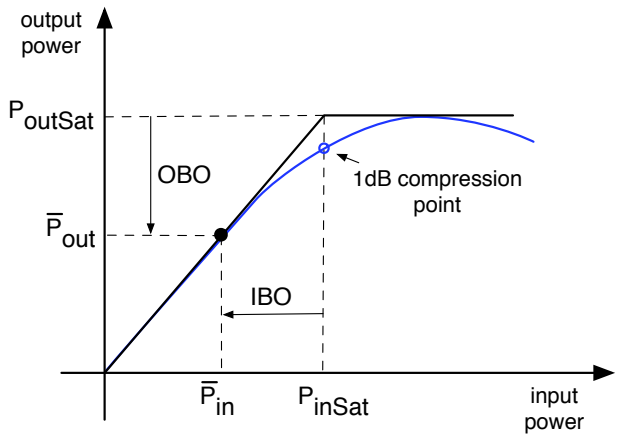


Figure 2: Typical AM/AM characteristic of an HPA.

The amplifier output power is always limited to some value P_{outSat}, called *saturation*. The point of the AM/AM characteristic that deviates 1dB from an extension of its linear part is known as the *1dB compression point* [54, 70], and it is commonly used to define the linear region of the HPA.

Both points, the saturation and the 1dB compression point, allow the definition of a simple *soft-limiter* model to characterize the nonlinear behavior of an HPA. The soft-limiter's AM/AM, $F(\rho)$, and AM/PM, $\Phi(\rho)$, characteristics are:

$$F(\rho) = \begin{cases} \rho \cdot \sqrt{G_{HPA}} & , \ \rho \leqslant A \\ A \cdot \sqrt{G_{HPA}} & , \ \rho > A \end{cases} , \tag{1.5}$$

$$\Phi(\rho) = 0 , \tag{1.6}$$

where ρ is the amplitude of the input signal, G_{HPA} is the power gain of the HPA in the linear region[2] and A is specified by the 1dB compression point, i.e. $A = \sqrt{P_{inSat}}$ with P_{inSat} being the HPA input saturation power. This model for the HPA nonlinearity is considered to be adequate when the HPA is linearized by a suitable pre-distorter [67, 108], and it will be the one used throughout the thesis.

To avoid the *clipping* of a signal having a high peak power (with consequent undesirable spectral spreading and in-band distortion) we must guarantee that the HPA is operated in its linear region by decreasing the average power of the input signal, so that the peak power of the input signal be less than P_{inSat}. This is called *input back-off* (IBO) and results in a proportional output back-off (OBO), as shown in figure 2, with both values being measured in decibels

$$IBO = 10 \log_{10}\left(\frac{P_{inSat}}{\overline{P}_{in}}\right) \quad (dB) , \tag{1.7}$$

$$OBO = 10 \log_{10}\left(\frac{P_{outSat}}{\overline{P}_{out}}\right) \quad (dB) , \tag{1.8}$$

where $\overline{P}_{in}$ and $\overline{P}_{out}$ are the average input and output of the HPA.

For the soft-limiter model defined by (1.5) and (1.6) both measures are equal, and so henceforth we will use only the generic designation of back-off (BO).

HPA efficiency

Linearity has costs in terms of the efficiency of a power amplifier. In fact there are three main classes of linear amplifiers: A, AB and B [59, 67, 94]. The classes are defined according to the biasing used, which corresponds to different conduction angles, i.e. several degrees of linearity. Class A amplifiers conduct for the entire cycle of the input signal, that is, the conduction angle is 2π and so they are the most linear. However, they are also the least efficient with the maximum power efficiency reaching only 50%. A higher conduction angle means better linearity and worse efficiency. Class B is characterized by a conduction angle of π, and class AB for an angle in $]\pi, 2\pi[$.

Worse still, a high PAPR dramatically reduces the power efficiency of the HPA. A rule of thumb is that for each 1 dB increase in required power amplifier

2 Hereafter, we will assume $G_{HPA} = 1$ without any loss of generalization.

headroom, a corresponding 10 to 15% increase in PA current drain occurs, leading to a corresponding reduction in battery life. In addition to inefficiency in terms of power, the coverage range is reduced, and the cost of the HPA is higher than would be warranted by the average power requirements. Figure 3 presents the theoretical efficiency limits of linear amplifiers [74] as function of the input signal PAPR.

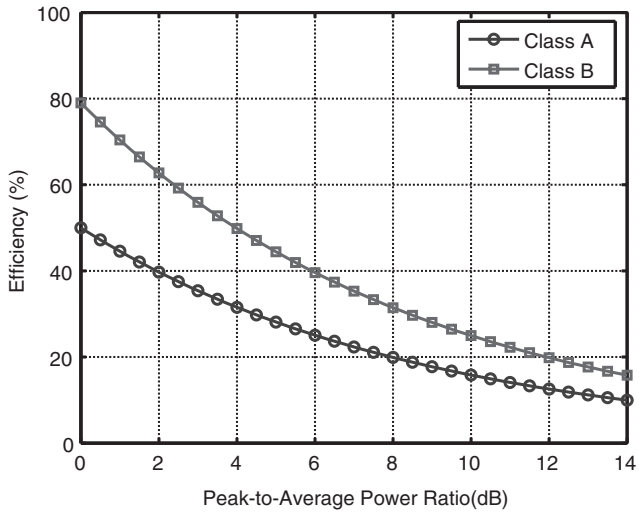


Figure 3: Theoretical efficiency limits of linear amplifiers [74].

Reducing PAPR of strictly band-limited *single-carrier* signals is therefore a matter of great concern in *bandwidth limited channels*.

However, in practice, the HPA should at least be able to handle $PAPR_{Const}$, i.e. if the multirate pulse shaping system (L up-sampler followed by filter $h[n]$) has unitary power gain, $G_{RRC} = 1$, then maximum amplitude symbols (with amplitude denoted as A_{max}) at the output of modulator (figure 1) should suffer no distortion if they are fed directly at the HPA input. The main goal of any PAPR reduction technique for SC transmission is, therefore, to cancel the undesirable $PAPR_{RRC}$ contribution, eliminating the need to back-off from HPA saturation, which can be expressed as

$$BO = 10\log_{10}\left(\frac{\max|x[n]|^2}{A_{max}^2 G_{RRC}}\right) \quad (dB)\,, \tag{1.9}$$

and this is the main objective of this work.

1.2.2 *DAC*

Although less relevant, the influence of a high PAPR on the requirements of the transmitter's DAC and the receiver's ADC should also be taken in account. High resolution DAC and ADC must be used to deal with large peaks while maintaining adequate signal resolution to overcome noise generated in the digital-to-analog conversion process. Hence, reducing the PAPR makes the design of these converters easier, thereby reducing their cost.

1.3 RELATED WORK

As shown previously, reducing PAPR has always been of major importance in any communication system, mainly due to linearity requirements put on the transmitter's HPA being related to a considerable decrease of system's power efficiency.

HPA's linearization by means of feedback or pre-distortion techniques [59] has always been a mean of operating the amplifier closer to saturation, and therefore improving its efficiency. But that doesn't solve the loss in power efficiency from high PAPR, since the amplifier must still operate with a large amount of back-off to avoid clipping of the transmitted signal.

There is a vast number of publications on the peak power problem, although until recently only a few have considered the SC case. However, the large number of patents on the subject underlines the importance of this theme for the industry.

One way to provide a signal with a lower PAPR to the DAC+HPA block consists of minimizing the $\text{PAPR}_{\text{Const}}$ contribution. As already mentioned, in the recent DVB-S2 satellite standard [19] a new set of amplitude phase shift keying (APSK) constellations has been adopted [15, 16] which combines a high spectral efficiency with a low PAPR compared with M-QAM constellations providing the same spectral efficiency.

Another solution for $\text{PAPR}_{\text{Const}}$ minimization is to avoid the transmission of critical sequences of symbols which may originate large peaks, e.g. a sequence of successive symbols having large magnitude and similar phase or whose phases change from one symbol to the next approximately[3] 180^{o}. This may be accomplished by systematically picking consecutive symbols from different constellations (usually non-square, for better performance) related to each other through constellation-phase rotation [81, 97].

These constellation-oriented techniques have a limited peak power control capability, however. As an alternative, a sequence-oriented approach may be followed that avoids the problematic sequence of modulated symbols through the use of coding [64, 92, 103] or more complex *trellis shaping* coding schemes

3 Both these sequences combine constructively with the pulse shaping impulse response to produce large peaks.

[13, 104]. The major drawback of these solutions is the slower the information rate. In addition, even though some trellis solutions may significantly decrease the PAPR, they are usually complex, with a high computational overhead.

Although $PAPR_{Const}$ is relevant, a major contribution to PAPR comes from Nyquist pulse shaping for limiting signal bandwidth without ISI. Some solutions to decrease $PAPR_{RRC}$ by optimizing the Nyquist pulse shape have therefore been proposed [11, 88].

We can divide PAPR reduction techniques into two main groups, *post-* and *pre-* pulse shaping filtering techniques, with all the above described methods being included in the latter category. But the best known peak power control methods always imply some sort of distortion on the transmitted signal.

One of the simplest post-filtering methods of decreasing peak power of bandwidth limited signal $x[n]$ is to deliberately *clip* the signal before amplification [63, 67]. This generates serious spectral spreading, however, which is the main reason to avoid operating HPA near saturation and, therefore, the motive for trying to reduce signal's PAPR. Nonetheless, many post-filtering techniques use the concept of error shaping to decrease the signal's PAPR [48, 85, 109]. Basically, the error of hard-clipping is bandwidth limited by a proper shaping

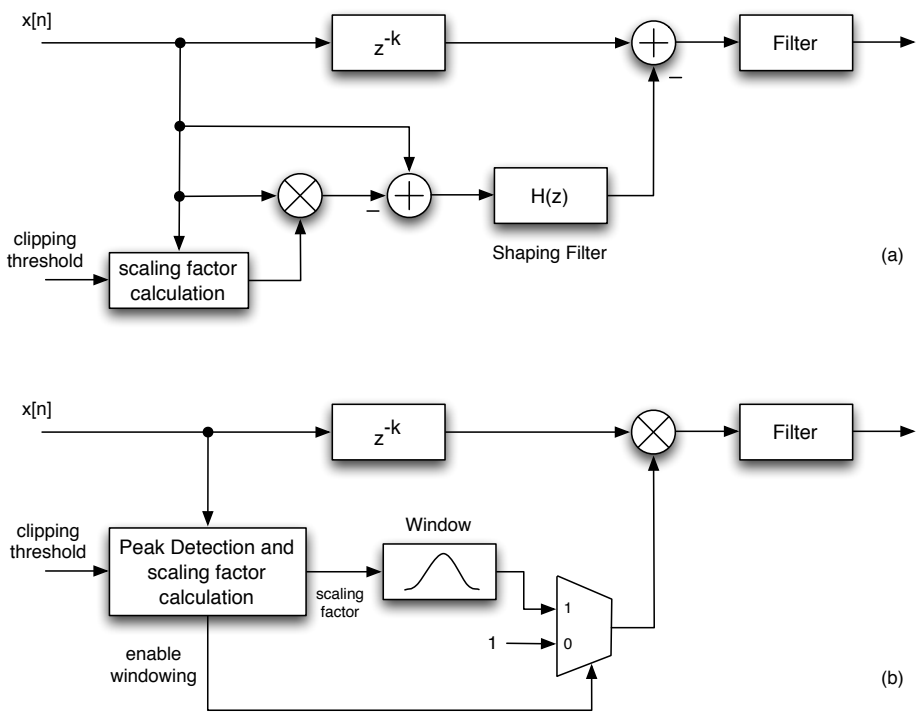


Figure 4: Post filtering peak power reduction techniques: (a) error shaping, (b) windowing.

filter and then subtracted from a delayed version of the original undistorted signal, as shown in figure 4(a). Another common technique is the concept of *peak windowing* [79, 109], shown in figure 4(b), that basically consists of detecting undesirable peaks of $x[n]$, with the signal being multiplied in those locations by a proper scaled window (e.g. Hamming, Hanning, Kaiser, etc.) of limited bandwidth. But multiplication in the time domain corresponds to a convolution in the frequency domain and spectral spreading results. In order to put the spectrum back within allocated limits, the windowing step is usually followed by a new filtering stage. Methods that combine the concepts of error shaping and windowing can also be found [40, 102]. Unfortunately, the final filtering stage in post-filtering techniques, required to guarantee bandwidth limitation, can cause PAPR regrowth. Most techniques solve this by using a cascade of these blocks, based on the premise that peak regrowth does not result in peaks as large as the original signal peaks. The major drawback is that multiple stages can introduce an unacceptable amount of latency into the communications system.

Making signal adjustments prior to filtering has the advantage of not affecting bandwidth of the signal $x[n]$. These adjustments must take into account the RRC impulse response $h[n]$ in order to also prevent the peak regrowth problem. Otherwise the efficiency gain is limited. Another pre-filtering technique is data *Magnitude Modulation* (MM). This consists of adjusting the amplitude of each data pulse $s[n]$ prior to RRC filtering. This concept, proposed in the nineties by Miller *et al.* [74, 75], and later developed by Tomlinson *et al.* [4, 108] who proposed an efficient *look-up table* (LUT) implementation, is the root of the work presented in this thesis.

A final reference should be made to some recent techniques [55, 56] that use the concept of *polyphase filtering* [3, 44, 76] to predict RRC response to the signal $x[n]$ when performing clipping of a signal. They are important because, at the same time, we have independently developed a similar strategy, giving rise to a new efficient *polyphase magnitude modulation* technique [35–38].

1.4 RESEARCH CONTRIBUTIONS AND ORGANIZATION OF THE THESIS

This thesis discusses the problem of controlling the envelope's power peak of SC modulated signals band-limited by Nyquist pulse shaping filters when using M-ary constellations of high order and filters with strict roll-offs.

Following the original work by Miller *et al.* [74, 75] and the later work of Tomlinson *et al.* [4, 108] we have concentrated on the concept of magnitude modulation (MM). This thesis contributes highly efficient algorithms for the control of the envelope's power peak through the use of MM, and also provides an analytical model for statistically characterizing the MM procedure.

1.4.1 *Research contributions*

The main contributions are discussed in the following chapters:

– Chapter 3

The LUT-MM approach, first proposed by Tomlinson *et al.* [4, 108] for QPSK modulation formats, is improved and extended to constellations with dimension up to $M \leqslant 16$ [32, 34, 38]. In order to cope with the additional complexity, new clipping limits are proposed and the method's ability to avoid phase modulation is improved [34]. A detailed analysis of the different approaches to compute the magnitude modulation tables is performed. We show that LUT's computation complexity and storage requirements can be considerably reduced by exploring the constellation's symmetry and the RRC's linear phase characteristic [32].

– Chapter 4

Building on the *polyphase decomposition* of the *finite impulse response* (FIR) of the pulse shaping filter, a new LUT-less approach of the magnitude modulation's concept is developed. We present a novel, highly efficient *Multistage Polyphase Magnitude Modulation* (MPMM) scheme [35–38] for decreasing the PAPR of an SC modulated signal, bandwidth limited by a *finite impulse response* (FIR) filter.

The problem related to an increased sensitivity to noise due to MM is also addressed. Two techniques are proposed for improving the detection of the MPMM signal, namely: detection based on the MM estimated average constellation (EAC) [35, 36] and detection with estimation of MPMM factors in the receiver.

– Chapter 5

A mathematical model based on parameterized beta distributions is also proposed for characterizing the statistical distributions of the MM factors for different SC constellations [38, 39]. Building on the proposed model, we carry out an analytical performance analysis of the MM technique in terms of the power of the transmitted signal, the *error vector magnitude* (EVM) and the *symbol error rate* (SER) experienced in a noiseless transmission, and on the *block error rate* (BLER) when channel coding is used. We present expressions for the analytical computation of the EVM and bounds on the SER, for the cases of MM's blind reception and detection based on the MM EAC, as well as bounds on the BLER performance of the MM scheme over an AWGN channel when using *forward error correction* (FEC) with linear block coding.

– Appendix B

Outside the main scope of the thesis, but included in related channel coding research activities, we propose encoding and decoding efficient hardware ar-

chitectures [30, 31, 33] for the *low-density parity-check* codes (LDPC) used by satellite's digital video broadcast DVB-S2 standard [19].

1.4.2 *Thesis organization*

This thesis is composed of 6 chapters and 2 appendices. The contents of chapters 3, 4 and 5 and appendix B have just been summarised.

Chapter 2 covers the basic concepts of error control coding. Channel coding plays a crucial role in improving communications reliability. The system's FEC capabilities are explored in chapters 3 and 4 in order to offset the increased noise sensitivity resulting from MM, which is why some reference must be made to this subject. In addition, the derived bounds, under hard- and soft-decoding, presented in this section, allow us to access the performance of MM techniques analytically through comparison with the matched ones that will be deduced in chapter 5 for the case where peak power control is used. Some attention is also devoted to the special case of the LDPC codes that were used in this work to evaluate BER performance of the peak power control techniques devised, and that were also the subject of some parallel research that is presented in appendix B.

Given the fact that a major contribution to the PAPR problem comes from Nyquist pulse shaping filters, usually achieved with RRC filters, it makes sense to briefly discuss their characteristics. This is done in appendix A which also contains the proof that

$$\mathrm{PAPR} = \mathrm{PAPR}_{\mathrm{Const}} + \mathrm{PAPR}_{\mathrm{RRC}} \ . \tag{1.10}$$

Finally, chapter 6 summarizes the main conclusions and provides some of the bearing lines of future research.

2

ERROR CORRECTING CODING

The seeds of *Error Correcting Coding* were sown by Claude Shannon in his 1948 landmark paper [99]. Shannon demonstrated that if the information is properly encoded errors induced by a noisy channel or storage medium can be reduced to any desired level without sacrificing the rate of information transmission or storage, as long as the information rate is less than the capacity of the channel. For the next fifty years the core objective of research was to find practical coding schemes that could come near to the *Shannon limit*, i.e. that could approach channel capacity on channels such as the *additive white Gaussian noise* (AWGN), the *binary symmetric channel* (BSC) and the *Rayleigh fading channel*, used to model typical transmission and storage noise disturbances. In the past decade, with the advent of *turbo codes* [6] and the rebirth of *low-density parity-check* codes (LDPC) [20, 69], this goal has been partly accomplished and these developments have helped to achieve the reliability required by today's high speed digital systems.

Coding for error control is nowadays an integral part of modern communication and digital storage systems. This chapter introduces the necessary coding notions. We focus on *forward error correction* (FEC) codes that are used in one-way channels (such as compact disk recordings, deep-space communication and digital video broadcasting systems, etc.), wherein the receiver must detect and correct errors without the use of a feedback channel (for retransmission requests). The most common types of codes are presented. The class of binary *linear block codes* is emphasized, however, with their performance being evaluated under *soft-* and *hard-decision* decoding for the AWGN channel. Finally, the LDPC family of capacity-approaching linear block codes, and the principles of *iterative soft-decoding* are described.

2.1 ERROR CORRECTING CODING: BASIC CONCEPTS

All error correcting codes are based on the same basic principle: redundancy is added to information in order to correct any errors that may occur in the storage or transmission process. Shannon [99] proved that in an AWGN channel, through proper encoding, the probability of information bit error, i.e. the *bit error rate* (BER), can be reduced to any desired level as long as the information transmission rate R_b [bits/s] does not exceed the *channel capacity* C:

$$C = W \log_2(1 + \frac{S}{N}) \ \text{[bits/s]} \ . \tag{2.1}$$

where S/N is the channel's *signal-to-noise ratio* (SNR), and W [Hz] the bandwidth available for transmission.

Typically, the encoder maps the sequence of bits at the rate R_b to a sequence of symbols at a baud rate R_s [Bd]. The *code rate* r_c is therefore,

$$r_c = \frac{R_b}{R_s} \text{ [bits/symbol]} . \tag{2.2}$$

According to Nyquist theory [87], the minimum bandwidth B required to transmit information at a rate R_s (also called the *Shannon bandwidth*) is

$$B = \frac{R_s}{2} . \tag{2.3}$$

This bandwidth is ideal and it is not achievable since it assumes an ideal bandwidth limiting filter. As stated in chapter 1, in practice, to avoid intersymbol interference we use a Nyquist filter with a roll-off factor α and so B relates to the bandwidth available for transmission W as follows,

$$W = B(1+\alpha) . \tag{2.4}$$

The *spectral efficiency* η [bits/s/Hz] of the coding scheme is thus evaluated as

$$\eta = \frac{R_b}{W} = \frac{2r_c}{1+\alpha} \text{ [bits/s/Hz]} . \tag{2.5}$$

As $B \to W$, the spectral efficiency η approaches $2r_c$, called the *nominal spectral efficiency*. Assuming ideal transmission, i.e. $B = W$, and so denoting $2r_c$ by η, according to Shannon's result we can establish different limits on each of the parameters R_b, η and the S/N, given the restrictions on either of the other two. Therefore using (2.1), and as stated before, the rate for reliable transmission of information is upper-bounded by

$$R_b < W \log_2(1+\frac{S}{N}) . \qquad (\Rightarrow \textit{ Shannon limit on } R_b) \tag{2.6}$$

In the same way, the spectral efficiency is upper-bounded by

$$\eta < \log_2(1+\frac{S}{N}) , \qquad (\Rightarrow \textit{ Shannon limit on } \eta) \tag{2.7}$$

or, given a spectral efficiency η, the SNR needed for reliable transmission is lower-bounded by

$$\frac{S}{N} > 2^{\eta} - 1 . \qquad (\Rightarrow \textit{ Shannon limit on SNR}) \tag{2.8}$$

These bounds define two different scenarios, the so-called *power-limited* ($\eta \leqslant 2r_c \leqslant 2$ [bits/s/Hz]) and *bandwidth-limited* ($\eta \geqslant 2$ [bits/s/Hz]) regimes, which require quite distinct coding schemes [14, 66].

Most work on channel coding has focused on binary codes whose code rate $r_c < 1$ and so $\eta < 2$. Hence, taking the transmitted power as a restriction (power-limited regime) and referring to (2.6), to maintain the same information rate R_b as an uncoded system, these coding schemes require *bandwidth expansion* by a factor $1/r_c$.

In fact, it was believed in the early 70s that *coding gain*[1] could be achieved only through bandwidth expansion and that coding could serve no useful purpose at spectral efficiencies $\eta \geqslant 2$. In communication applications with restricted bandwidth (bandwidth-limited regime), where large modulation alphabets were needed to achieve high spectral efficiencies, design emphasis was almost exclusively on constructing large signal sets in two-dimension Euclidean space that had the highest possible minimum Euclidean distance between signal points, given certain constraints on average and/or peak signal energy. This was obviously a misleading thought since coding gain can be obtained through the serial concatenation of error correcting coding and higher-order modulation. However, higher efficiency and coding gain can be achieved by considering coding and modulation as a single entity, as suggested by Massey in 1974. The joint design of codes and signal constellations is known as *coded modulation*, and the practical breakthrough in this field came with Ungerboeck's invention of *trellis-coded modulation* and the *multi-level coded modulation* devised by Imai and Hirakawa [66, 78].

2.1.1 *Power-limited regime: the Ultimate Shannon Limit*

As mentioned above, most research in coding theory has concentrated on designing good codes and efficient decoding algorithms for binary-input channels. A typical figure-of-merit used to evaluate the performance of a binary coding system at a given BER, over an AWGN channel with white spectral noise density $N_0/2$ [W/Hz], is the ratio between the average received energy per information bit E_b [J] and N_0, i.e. E_b/N_0. This ratio relates to the SNR of the channel and it is usually called the *signal noise ratio per information bit* [87]. E_b can be expressed as:

$$E_b = \int_0^{T_b} S \, dt = ST_b = \frac{S}{R_b} \, , \tag{2.9}$$

where T_b [s] is the time duration of one information bit and S [W] is the average transmitted power.

1 The formal definition of *coding gain* is presented in sub-section 2.1.1.

The *noise equivalent bandwidth* [10] equals the Shannon bandwidth[2] B and the total noise power is therefore:

$$N = \int_{-B}^{B} \frac{N_0}{2}\, df = N_0 B \,. \tag{2.10}$$

Hence, we can replace S/N in (2.6) to (2.8) by:

$$\frac{S}{N} = \frac{E_b R_b}{N_0 B} = \eta \frac{E_b}{N_0} \,, \tag{2.11}$$

where the last equality holds when assuming ideal transmission, i.e. $W = B$.

A superficial analysis of Shannon's equations (2.1) and (2.6) may lead to the misleading conclusion that, for a certain channel's SNR, the channel capacity C, and therefore R_b, can be as large as desired, provided the bandwidth W is increased. This is the same as saying that, given the desired information rate R_b, by allowing $W \to \infty$ (i.e. making the spectral efficiency $\eta \to 0$) the required E_b/N_0 will tend to zero, which is not true.

By replacing (2.11) in (2.8), the Shannon limit on SNR can be written as:

$$\frac{E_b}{N_0} > \frac{2^\eta - 1}{\eta} \,. \tag{2.12}$$

Then, as $\eta \to 0$, the lower bound on E_b/N_0 is:

$$\frac{E_b}{N_0} > \lim_{\eta \to 0} \frac{2^\eta - 1}{\eta} = \ln 2 = -1.59 \text{ [dB]} \,. \tag{2.13}$$

which means that even for an infinite bandwidth, E_b/N_0 has an absolute minimum value required for an error-free transmission. This value is usually called the *ultimate Shannon limit*.

Similarly, it can be concluded that the channel capacity C, in the limit as $W \to \infty$, is, in fact, upper-bounded by

$$\lim_{B \to \infty} C = 1.443\, R_b\, \frac{E_b}{N_0} \,. \tag{2.14}$$

The ultimate Shannon limit establishes a ceiling for the amount of *gain* that can be achieved by any error correcting scheme. This gain is usually measured as the difference in E_b/N_0 [dB], required to guarantee the same value of BER, between the coded system and an uncoded system with the same information rate R_b, and is referred to as the *coding gain* of the error correcting scheme.

2 The area under the frequency response curve of the Nyquist filter with bandwidth $B(1+\alpha)$, is the same as that of the ideal filter of bandwidth B [10].

2.1.2 *Block Codes and Convolutional Codes*

Depending on how redundancy can be added to messages, there are two important classes of codes: *block* and *convolutional* [7, 66].

In general, we define block codes over an arbitrary finite alphabet $\mathcal{X}$, say an alphabet with q symbols $\{0, 1, 2, \cdots, q-1\}$. Block codes process the information on a block-by-block basis, i.e. the encoder divides the information sequence into blocks of k information symbols each[3], which it transforms independently into blocks of n symbols (with $n > k$) of the same alphabet $\mathcal{X}$. Each k-tuple block of information symbols,

$$\mathbf{m} = [m_0, m_1, \cdots, m_{k-1}] \quad \text{with} \quad m_i \in \mathcal{X} \tag{2.15}$$

is called a *message*. Hence, there is a total of q^k different k-tuple messages that the block encoder maps to q^k distinct n-tuple vectors, called *codewords*, with each codeword being represented as,

$$\mathbf{c} = [c_0, c_1, \cdots, c_{n-1}] \quad \text{with} \quad c_i \in \mathcal{X}\,. \tag{2.16}$$

This set of q^k codewords of length n is called an (n,k) *block code*. The code rate as defined in (2.2) is thus given by[4]

$$r_c = \frac{k}{n}\log_2 q \ \text{ [bits/symbol]}\,. \tag{2.17}$$

The code is said to be *systematic* when the message $\mathbf{m}$ can be found explicitly and unchanged in the codeword $\mathbf{c}$. That is, the block encoder maps each message into a codeword that has k elements, say at coordinates $i_0, i_1, \cdots, i_{k-1}$, that equal the k symbols of the message, i.e. $c_{i_0} = m_0, \cdots, c_{i_{k-1}} = m_{k-1}$. In this case the remaining $n-k$ symbols $\mathbf{p}$ that are appended to the message are called *parity symbols*, and each codeword is usually represented as

$$\mathbf{c} = [\mathbf{m}\,|\mathbf{p}] = [m_0, \cdots, m_{k-1}, p_0, \cdots, p_{n-k-1}] \quad \text{with} \quad m_i, p_i \in \mathcal{X}\,, \tag{2.18}$$

where without loss of generality, a systematic encoder is assumed *on the left*[5].

---

3 In a digital transmission the input source data is usually taken to be binary ($q = 2$). This should not place any restriction on using non-binary q-ary block codes ($q > 2$), since it is a trivial matter to represent binary source data in terms of a q-ary alphabet, especially if q is a power of 2, as it usually is in practice.

4 Another popular definition of *code rate* of an (n,k) block code, defined over a q-ary alphabet $\mathcal{X}$, that it is usually found in literature, is $r_c = k/n$. This rate is dimensionless and $r_c \leqslant 1$. Both definitions agree for binary block codes, i.e. $q = 2$.

5 It means that the message $\mathbf{m}$ appears in the first k positions (on the *left*) of codeword $\mathbf{c}$.

Because a block encoder treats each message of information symbols independently of the others, the encoder is memoryless and can be implemented with a combinational logic circuit.

A convolutional code, on the other hand, is generated by passing the information sequence to be transmitted through a linear finite-state shift register (sequential logic circuit) with *memory order* L, called the *constraint length* of the convolutional code. Instead of blocks the convolutional encoder takes a non-ending sequence of information symbols arranged in k-symbol segments called *dataframes*, and outputs a sequence of code symbols arranged in n-symbol segments called *codeframes*. Each generated codeframe depends not only on the corresponding k-symbol dataframe at the same time unit but also on L previous dataframes.

Although both types of codes have found practical applications, historically, convolutional codes have been preferred, apparently because of the availability of the soft-decision *Viterbi decoding algorithm* [7, 66], and the belief for many years that block codes could not be efficiently decoded with soft-decisions. But the rediscovery of LDPC codes in the mid 90s [20, 69] and the developments in the theory and design of soft-decision decoding algorithms for linear block codes helped to dispel this belief.

2.2 LINEAR BLOCK CODES

As mentioned above, given an (n, k) block code, defined over a q-ary alphabet $\mathcal{X}$, encoding is basically a bijective assignment between the q^k different k-tuple messages and q^k n-tuple codewords of symbols of $\mathcal{X}$. When k and n are small this assignment can be done by means of a table, but when they are large efficient encoding and decoding mechanisms must be found. If algebraic operations could be used this would greatly simplify the encoding and decoding procedures.

Algebraic coding dominated the first few decades of the research in error correcting coding. Particularly, the Galois' theory of *finite fields* and *vector spaces* defined over those fields [7, 66], is on the basis of one the most important classes of error correcting codes, the *linear block codes* [101].

In general, a field F consists of a set of elements that has two arithmetic operations defined in them, i.e. addition and multiplication, that satisfy a certain number of axioms [7, 26]. A finite field with q elements, usually denominated by *Galois Field* GF(q), can only be constructed whenever q is a prime number, or the power of a prime number, and has the elements $\{0, 1, 2, \cdots, q-1\}$. Addition and multiplication defined in GF(q) are modulo-q operations whenever q is a prime number. When q is the power of a prime number p, i.e. $q = p^m$ with m an integer, the field GF(p^m) is called the *extension field* of GF(p) and the addition and multiplication operations in the extension field are based on modulo-p arithmetic.

The simplest and also the most common GF used in coding theory is the binary field GF(2), with $\mathcal{X} = \{0, 1\}$, which has the obvious connection to the logical symbols "0" and "1" called bits. In fact, we easily conclude that modulo-2 addition and multiplication in GF(2) are simply the *boolean* operations *EXOR* and *AND* performed on bits.

The set of all n-tuple vectors over GF(q) is a *vector space* [7, 26] of *dimension* n, denoted as GF(q)^n. An (n, k) block code $\mathcal{C}$ is therefore a non-empty subset of q^k different vectors of GF(q)^n.

A block code $\mathcal{C}$ is said to be *linear* if and only if the linear combination of any number of codewords is also a codeword, i.e. suppose $\mathbf{c}_i$ and $\mathbf{c}_j$ are any two codewords of $\mathcal{C}$, and let α_1 and α_2 be any two elements from GF(q), then the code is linear if and only if $\alpha_1\mathbf{c}_i + \alpha_2\mathbf{c}_j$ is a codeword as well. Hence, an (n, k) block code is linear if and only if the set of q^k codewords forms a k-dimensional *subspace* of GF(q)^n. An obvious conclusion to draw from this definition is that in a linear code the all-zero n-tuple is always a codeword.

2.2.1 *Generator and Parity-Check Matrices*

Because an (n, k) linear code $\mathcal{C}$ is a k-dimensional subspace of GF(q)^n, it is possible to find in $\mathcal{C}$ a *basis* of k linearly independent codewords, say $\mathbf{g}_0, \mathbf{g}_1, \cdots \mathbf{g}_{k-1}$, such that any codeword $\mathbf{c}$ in $\mathcal{C}$ is a linear combination of these k codewords, i.e.

$$\mathbf{c} = m_0\mathbf{g}_0 + m_1\mathbf{g}_1 + \cdots + m_{k-1}\mathbf{g}_{k-1} \,, \tag{2.19}$$

where m_i can be any element of GF(q), for $i = 0, \cdots k-1$. Because there are q^k distinct vectors $[m_0\ m_1\ \cdots\ m_{k-1}]$ over GF(q) that can be chosen, each of these vectors can be seen as each of the k-tuple messages $\mathbf{m}$ to be encoded.

The linear equation (2.19) can thus represented in a matrix form as

$$\mathbf{c} = [m_0\ m_1\ \cdots\ m_{k-1}] \cdot \begin{bmatrix} \mathbf{g}_0 \\ \mathbf{g}_1 \\ \vdots \\ \mathbf{g}_{k-1} \end{bmatrix} = \mathbf{m} \cdot \mathbf{G} \,, \tag{2.20}$$

where $\mathbf{G}$ is called a *generator matrix* of the code, because the rows of $\mathbf{G}$ *generate*, i.e. span, the (n, k) linear code $\mathcal{C}$.

Matrix $\mathbf{G}$ not only completely characterizes the code $\mathcal{C}$, it also defines the bijective assignment between the messages $\mathbf{m}$ and the codewords $\mathbf{c}$ of $\mathcal{C}$. However, since the set of basis codewords used to span $\mathcal{C}$ is not unique, we conclude that $\mathbf{G}$ is not unique, either. In fact, any algebraic operations or permutations performed on the rows of a generator matrix $\mathbf{G}$ give rise to a different generator

matrix **G'** that spans $\mathcal{C}$ as well[6]. Matrix **G'** only defines a different assignment between the messages and the set of codewords of $\mathcal{C}$.

However, permutations on the columns of **G** give rise to a generator matrix of a different linear code $\mathcal{C}'$, i.e. a code with a different set of codewords. Nevertheless, as it will be clear in next sub-section 2.2.2, both $\mathcal{C}$ and $\mathcal{C}'$ have the same error correcting capabilities and so they are said to be *equivalent codes*.

Any (n, k) linear code can thus be made systematic, by reducing its generator matrix to a *systematic* (or *canonic*) *form* as

$$\mathbf{G} = [\mathbf{I}_k | \mathbf{P}] = \left[\begin{array}{ccccc|cccc} 1 & 0 & 0 & \cdots & 0 & p_{00} & p_{01} & \cdots & p_{0,n-k-1} \\ 0 & 1 & 0 & \cdots & 0 & p_{10} & p_{11} & \cdots & p_{1,n-k-1} \\ \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & 1 & p_{k-1,0} & p_{k-1,1} & \cdots & p_{k-1,n-k-1} \end{array}\right] \tag{2.21}$$

where $\mathbf{I}_k$ is the $k \times k$ identity matrix and **P** is a $k \times (n-k)$ matrix that determines the $n-k$ redundant symbols. The particular systematic form (2.21), with $\mathbf{I}_k$ on the leftmost k columns of **G**, maps each message **m** into codewords having the format (2.18). Redundant (check) symbols are thus obtained through a set of $n-k$ linear equations,

$$\mathbf{p} = \mathbf{m} \cdot \mathbf{P} \,. \tag{2.22}$$

Because $\mathcal{C}$ is a k-dimensional sub-space of $\mathrm{GF}(q)^n$, it has an orthogonal complement $\mathcal{C}^\perp$, which is the set of all vectors orthogonal to $\mathcal{C}$. The orthogonal complement is also a sub-space of $\mathrm{GF}(q)^n$ (the null-space of $\mathcal{C}$) and it has dimension $n-k$. Hence, $\mathcal{C}^\perp$ itself is an $(n, n-k)$ linear block code and it is called the dual code of $\mathcal{C}$.

Let **H** be a generator matrix for $\mathcal{C}^\perp$. Because the set of rows of **H** is a basis of $\mathcal{C}^\perp$ and since all vectors in $\mathcal{C}^\perp$ are orthogonal to $\mathcal{C}$, matrix **H** can also be used to describe the code $\mathcal{C}$. In fact, an n-tuple **c** is a codeword in $\mathcal{C}$ if and only if

$$\mathbf{c} \cdot \mathbf{H}^T = \mathbf{0} \,, \tag{2.23}$$

where **0** is an $(n-k)$-tuple vector of zeros. As **H** provides a way for checking whether a word is a codeword of $\mathcal{C}$, it is called the *parity-check matrix* of the code $\mathcal{C}$. Clearly, it results from (2.23) that

$$\mathbf{G} \cdot \mathbf{H}^T = \mathbf{0}_{k \times (n-k)} \,. \tag{2.24}$$

6 Given the code's linearity property, by performing arithmetic operations or permutations over the rows of **G**, we obtain a different ordered set of k linearly independent codewords that constitute **G'**.

The parity-check matrix $\mathbf{H}$ can also be reduced to a systematic form by elementary row operations. If the generator matrix of $\mathcal{C}$ has the systematic structure shown in (2.21), then, according to (2.24), the systematic form for $\mathbf{H}$ is

$$\mathbf{H} = \left[\mathbf{P}^T \middle| \mathbf{I}_{n-k} \right] . \tag{2.25}$$

2.2.2 *Error-Detecting and Error-Correcting Capabilities*

Any *full-rank* $k \times n$ matrix defined over GF(q) can be used as generator $\mathbf{G}$ of an (n, k) linear block code, or as parity-check matrix $\mathbf{H}$ of its $(n, n-k)$ dual code. However, not all possible codes (n, k) defined over $\mathrm{GF}(q)^n$ have good *error-detecting* and *error-correcting* capabilities. A "good" code should have its codewords as far apart as possible.

A measure of the difference between two codewords, say $\mathbf{c}_i$ and $\mathbf{c}_j$, is the *Hamming distance*, denoted as $d_H(\mathbf{c}_i, \mathbf{c}_j)$ and defined as the number of corresponding elements or positions in which $\mathbf{c}_i$ and $\mathbf{c}_j$ differ. The shortest distance between all distinct pairs of codewords in $\mathcal{C}$ is called the *minimum distance*, d_{min} of the code, i.e.

$$d_{min} = \min_{\forall \mathbf{c}_i, \mathbf{c}_j \in \mathcal{C}} \left\{ d_H(\mathbf{c}_i, \mathbf{c}_j) \middle| \mathbf{c}_i \neq \mathbf{c}_j \right\} . \tag{2.26}$$

Another important measure is the *Hamming weight* of a codeword $\mathbf{c}$, denoted as $w_H(\mathbf{c})$, which is the number of non-zero elements of the codeword. If the code is linear then it follows from the definition of modulo-q addition that

$$d_H(\mathbf{c}_i, \mathbf{c}_j) = w_H(\mathbf{c}_i + \mathbf{c}_j) . \tag{2.27}$$

Hence, the weight distribution of a linear code completely characterizes the distance properties of the code. Furthermore, the minimum distance of the code is exactly the weight of the non-zero codewords with smallest weight, i.e.

$$d_{min} = \min_{\mathbf{c} \in \mathcal{C} \setminus \{\mathbf{0}\}} \{ w_H(\mathbf{c}) \} . \tag{2.28}$$

Both d_{min} and the weight distribution of a linear code $\mathcal{C}$ specify the error-detecting and error-correcting capabilities of $\mathcal{C}$, and the performance under soft- and hard-decision decoding. Since any two distinct codewords of $\mathcal{C}$ differ at least d_{min} places, no error pattern of $d_{min} - 1$ or fewer errors occurring during transmission or storage can change one codeword into another. Therefore, $\mathcal{C}$ guarantees detection of all error patterns of $d_{min} - 1$ or fewer errors and so the code is said to be capable of detecting t_{det} errors with

$$t_{det} = d_{min} - 1 . \tag{2.29}$$

Although $\mathcal{C}$ cannot detect all error patterns with d_{min} or more errors, it can detect any error pattern that does not equal a codeword, i.e. all the $q^n - q^k$ received words that violate (2.23).

To determine the error-correction capability of $\mathcal{C}$, it is convenient to view its q^k codewords as points in an n-dimensional space. If each codeword is viewed as the centre of a sphere of radius (Hamming distance) t_{cor}, the largest value that t_{cor} may have without intersection or tangency of any pair of the q^k spheres is

$$t_{cor} = \left\lfloor \frac{d_{min} - 1}{2} \right\rfloor , \tag{2.30}$$

where $\lfloor x \rfloor$ denotes the largest integer contained in x. Hence, the linear code $\mathcal{C}$ is said to be capable of correcting all error patterns of t_{cor} or fewer errors.

If these q^k spheres span all vector space GF$(q)^n$, then $\mathcal{C}$ is called a *perfect code*. This means that every received word is no more than distance t_{cor} from one of the possible transmitted codewords and therefore the decoder is always able to take a decision on the most probable codeword that was sent. Because there are $\binom{n}{l}$ ways to choose l places where a point differs from a codeword of $\mathcal{C}$, and $q-1$ ways in which it can be different in each of those places, the number of points contained in each *Hamming sphere* of radius t_{cor} is

$$\sum_{l=0}^{t_{cor}} \binom{n}{l} (q-1)^l , \tag{2.31}$$

and so if $\mathcal{C}$ is perfect then

$$q^n = q^k \times \sum_{l=0}^{t_{cor}} \binom{n}{l} (q-1)^l . \tag{2.32}$$

By making use of the same Hamming sphere concept, we conclude that $\mathcal{C}$ can detect e_d errors and correct e_c (with $e_c \leqslant e_d$) if

$$e_d + e_c < d_{min} . \tag{2.33}$$

2.2.3 *Decoding: Syndrome's Method*

Because codewords of $\mathcal{C}$ are the only solutions of (2.23) we use an easy algebraic technique to check whether a received word is a valid codeword without having to search the entire set of q^k codewords for the received vector.

Equation (2.23) also allows us to conclude that the minimum distance of $\mathcal{C}$ is equal to the smallest number of columns of $\mathbf{H}$ that sum to zero. In fact,

the structure of the **H** matrix not only dictates the error-detecting and error-correcting capabilities of the code but it also influences the performance of *iterative algorithms* used for decoding long codes.

One of the "simplest" and most efficient methods of decoding an (n, k) linear block code $\mathcal{C}$ is based on the concept of *syndrome*. Let **c** be a codeword of $\mathcal{C}$ and $\mathbf{e} \in \mathrm{GF}(q)^n$ be an error occurring during transmission. The received word **r** is thus given by

$$\mathbf{r}=\mathbf{c}+\mathbf{e} \ . \tag{2.34}$$

When **r** is received, the decoder computes the following $(n-k)$-tuple

$$\mathbf{s} = \mathbf{r}\,\mathbf{H}^T \ , \tag{2.35}$$

which is called the *syndrome* of **r**. According to (2.23), $\mathbf{s} = \mathbf{0}$ if and only if **r** is a codeword, and so the presence of errors is only detected when $\mathbf{s} \neq \mathbf{0}$. In this case, the decoder can either act to locate the errors and correct them (FEC) or request a retransmission of **c** (ARQ).

Because **r** is the vector sum of **c** and **e**, it follows from (2.23) that

$$\mathbf{s} = (\mathbf{c}+\mathbf{e}) \cdot \mathbf{H}^T = \mathbf{c} \cdot \mathbf{H}^T + \mathbf{e} \cdot \mathbf{H}^T = \mathbf{0} + \mathbf{e} \cdot \mathbf{H}^T = \mathbf{e} \cdot \mathbf{H}^T \ . \tag{2.36}$$

Hence, the syndrome **s** computed from the received word **r** depends only on the error pattern **e** and not on the transmitted codeword **c**. Also we conclude that there are q^k received words that produce the same syndrome, since **c** can be any codeword of $\mathcal{C}$.

Decoding can thus be performed based on a table, called *standard array* [7, 66]. In this array, each row, called a *coset*, contains all the error patterns (received words) that produce the same syndrome. There are $q^{(n-k)}$ syndromes and so the standard array has dimensions $q^{(n-k)} \times q^k$.

The error pattern with the smallest Hamming weight in each coset is called the *coset leader*. If the coset leader $\mathbf{e_s}$ of a coset is unique then the code is able to correct the received word **r** that produces the corresponding coset's syndrome **s**. The decoded word $\mathbf{c'}$ is thus given by

$$\mathbf{c'} = \mathbf{r} + \mathbf{e_s} \ . \tag{2.37}$$

2.2.4 *Hamming and Golay codes*

In this section we describe two of the most important non-trivial perfect binary codes that were discovered in the early days of error-correcting coding: *Hamming* and *Golay* codes.

Other important codes discovered early on, like the *cyclic redundancy check* (CRC), the BCH and the Reed-Solomon (RS) codes [66, 78], belong to a special class of linear block codes known as *cyclic codes*, and also deserve to be mentioned, mainly because both are still being used in today's major commercial applications [14]. For instance, CRC codes are widely used for error detection in conjunction with ARQ protocols in computer communications. A considerable number of them are international standards for error detection in various contexts [61, 66]; BCH are used as outer codes in the concatenated code system adopted by the second generation of the *satellite's digital video broadcast standard* (DVB-S2) [19], and RS codes are used in the compact disc (CD) standard [52]. They are beyond the scope of the main subject of the thesis, however, so they are not dealt with here.

Hamming Codes

Hamming codes are an important family of binary perfect codes capable of correcting a single error. They were proposed in 1950 by Richard Hamming [41], a colleague of Shannon at Bell Labs.

The dimensions (n, k) of a Hamming code must therefore satisfy (2.32) with $q = 2$ and $t_{cor} = 1$, meaning that

$$2^n = 2^k \times \sum_{l=0}^{1} \binom{n}{l} (q-1)^l \;\Rightarrow\; n = 2^{n-k} - 1 \,. \tag{2.38}$$

For any positive integer $m \geqslant 2$, therefore, there is a Hamming code of dimensions $(n = 2^m - 1, k = 2^m - m - 1)$, where m is the number of parity-check bits of the code, i.e. $m = n - k$.

In order to correct a single error, according to (2.30) a linear code must have $d_{min} \geqslant 3$ and so the columns of its parity check matrix $\mathbf{H}$ must be all distinct. Because $\mathbf{H}$ has dimensions $(n-k) \times n$ there are $2^{n-k} - 1$ non-zero $(n-k)$-tuple vectors that can be chosen for columns of $\mathbf{H}$. Since a Hamming code has $n = 2^{n-k} - 1$ its parity-check matrix consists of all these distinct $2^{n-k} - 1$ non-zero vectors in its columns, and so its minimum distance is exactly $d_{min} = 3$.

The structure of a Hamming code's $\mathbf{H}$ matrix makes decoding very simple when the syndrome method is used. If the computed syndrome $\mathbf{s}$ of a received word $\mathbf{r}$ is non-zero, then $\mathbf{s}^T$ is a column of $\mathbf{H}$. Let i denote the index of the column of $\mathbf{H}$ which is equal to $\mathbf{s}^T$. The coset leader of the syndrome's coset is a binary vector which is all zeros except for a "1" in the i^{th}-position, i.e. the i^{th}-bit of $\mathbf{r}$ is the one considered to be wrong.

The weight distribution of an (n, k) Hamming code is known [66, 87]. The number of codewords of weight i, A_i, is simply the coefficient of z^i in the expansion of the following polynomial:

$$A(z) = \frac{1}{n+1}\left[(1+z)^n + n\,(1+z)^{(n-1)/2}\,(1-z)^{(n+1)/2}\right], \tag{2.39}$$

usually called the *weight enumerating polynomial* of the code.

Golay Codes

Shortly after the publication of Shannon's paper the Swiss mathematician Marcel Golay published a half-page paper [27] presenting two perfect codes: a binary linear $(23, 12)$ triple-error-correcting code having minimum distance $d_{min} = 7$; and ternary linear $(11, 6)$ double-error-correcting code having minimum distance $d_{min} = 5$.

Their beautiful algebraic structure led them to be studied by many coding theorists and mathematicians who published numerous papers on their structure and decoding. An important result was obtained in 1974 and 1975 by Tietäväinen and Van Lint respectively, who proved that no other non-trivial perfect linear codes exist apart from the Hamming and Golay codes,[7, 77].

Beautiful structure apart, the binary $(23, 12)$ Golay code and its *extension* [66], a binary $(24, 12)$ linear code[7] with $d_{min} = 7$, have been used in many real and important communication systems, like the U.S. space program[8].

The Golay codes should therefore probably be considered as the most remarkable of all algebraic codes.

The weight enumerating polynomial of the $(23, 12)$ Golay code is also known and it is given by:

$$A(z) = 1 + 253\left(1+z^9\right)z^7 + 506\left(1+z^7\right)z^8 + 1288\,(1+z)\,z^{11} + z^{23}\,. \tag{2.40}$$

2.3 PERFORMANCE BOUNDS OF LINEAR BLOCK CODES OVER THE AWGN CHANNEL

In this section, we derive some bounds on the probability of a decoding error for a linear binary block code over an AWGN memoryless channel with zero mean white noise spectral density $N_0/2$ [W/Hz]. We assume that coherent BPSK signaling is used for transmission and optimum reception is performed by *matched filtering* in the receiver demodulator (or the use of *correlation demodulator*) [10, 87].

7 The extended binary $(24, 12)$ Golay code is capable of correcting all error patterns of three or fewer errors and detecting all error patterns of four errors.

8 The $(24, 12)$ Golay served as the primary *Voyager* error-control system, providing clear colour pictures of Jupiter and Saturn between 1979 and 1981.

Let $\mathcal{C}$ be a binary (n, k) linear block code with minimum Hamming distance d_{min} and weight enumerating polynomial $A(z)$. Assume that each transmitted BPSK symbol has energy E_s and let $\mathbf{c} = [c_0, c_1, \cdots, c_{n-1}]$ be a codeword $\in \mathcal{C}$. For transmission, this codeword is mapped into a sequence of BPSK signals that, in vector form, is represented by a bipolar sequence[9] $\mathbf{s} = [s_0, s_1, \cdots, s_{n-1}]$, where for $0 \leqslant i < n$,

$$s_i = \sqrt{E_s}\,(2c_i - 1) = \begin{cases} -\sqrt{E_s} & \text{for } c_i = 0 \\ \sqrt{E_s} & \text{for } c_i = 1 \end{cases} . \tag{2.41}$$

Let $\mathbf{r} = [r_0, r_1, \cdots, r_{n-1}]$ denote the n sampled outputs of the matched filter. The AWGN channel is characterized as having a finite alphabet at the input and producing a continuous output that can take any real value. Hence, each received sampled output may be expressed as

$$r_i = s_i + n_i = \sqrt{E_s}\,(2c_i - 1) + n_i\,, \tag{2.42}$$

where n_i is a real random variable representing the AWGN noise added to the transmitted symbol s_i, and so, the received sample r_i conditioned on s_i (or c_i, since there is a one-to-one correspondence between c_i and s_i) is a Gaussian random variable whose conditional *probability density function* (*pdf*) is given by:

$$p(r_i\,|s_i\,) = p(r_i\,|c_i\,) = \frac{1}{\sqrt{\pi N_0}} e^{-\frac{(r_i - s_i)^2}{N_0}} . \tag{2.43}$$

Since the channel is considered to be memoryless, the noise components $\{n_i\}$ are uncorrelated Gaussian random variables and so statistically independent. As consequence, the elements of $\mathbf{r}$ conditioned on $\mathbf{s}$ being transmitted are also statistically independent Gaussian variables and the joint conditional *pdf* is simply

$$p(\mathbf{r}\,|\mathbf{s}\,) = p(\mathbf{r}\,|\mathbf{c}\,) = \prod_{i=0}^{n-1} p(r_i\,|s_i\,) = \frac{1}{(\pi N_0)^{n/2}} \exp\left[-\sum_{i=0}^{n-1} \frac{(r_i - s_i)^2}{N_0} \right] . \tag{2.44}$$

2.3.1 *Decoding: MAP and ML decision rules*

Based on the received sequence $\mathbf{r}$, the decoder must produce an estimate $\hat{\mathbf{c}}$ of the transmitted codeword $\mathbf{c}$. A *decoding error* occurs whenever $\hat{\mathbf{c}} \neq \mathbf{c}$, and so the

9 Henceforth, the symbol $\mathbf{s}$ will be used to represent a vector of modulated symbols, and therefore should not be confused with the syndrome $\mathbf{s} = \mathbf{r}\,\mathbf{H}^T$ defined in section 2.2.3.

conditional error probability of the decoder, given that $\mathbf{r}$ was received, is defined as

$$\Pr(\mathrm{E}|\mathbf{r}) \triangleq \Pr(\hat{\mathbf{c}} \neq \mathbf{c}|\mathbf{r}) \ . \tag{2.45}$$

An optimal decoding rule should therefore minimize the overall error probability among all possible received sequences $\mathbf{r}$, i.e. it should minimize

$$\Pr(\mathrm{E}) = \int_{\mathbf{r}} \Pr(\mathrm{E}|\mathbf{r})\, p(\mathbf{r})\, \mathrm{d}\mathbf{r} \ , \tag{2.46}$$

where $p(\mathbf{r})$ is the joint pdf of the received sequence $\mathbf{r}$. However, since $p(\mathbf{r}) \geqslant 0$, an optimum decoding rule that minimizes $\Pr(\mathrm{E})$ is the one that minimizes $\Pr(\mathrm{E}|\mathbf{r})$ for all $\mathbf{r}$.

Because minimizing $\Pr(\mathrm{E}|\mathbf{r})$ is equivalent to maximizing $\Pr(\hat{\mathbf{c}} = \mathbf{c}|\mathbf{r})$ the optimum decoding rule should select a codeword $\hat{\mathbf{c}} \in \mathcal{C}$ whose *a posteriori probability* is greater than for any of the remaining codewords, i.e. $\hat{\mathbf{c}}$ should be selected such that

$$\Pr(\hat{\mathbf{c}}|\mathbf{r}) \geqslant \Pr\big(\mathbf{c}_j|\mathbf{r}\big) \quad , \forall \mathbf{c}_j \in \mathcal{C} \text{ with } j = 0, \ldots, 2^k - 1 \ . \tag{2.47}$$

This decision criterion is called the *maximum a posteriori probability* (MAP) criterion. The a posteriori probabilities $\Pr\big(\mathbf{c}_j|\mathbf{r}\big)$ can be computed using Bayes' rule, i.e. according to

$$\Pr(\mathbf{c}|\mathbf{r}) = \frac{p(\mathbf{r}|\mathbf{c})\Pr(\mathbf{c})}{p(\mathbf{r})} = \frac{p(\mathbf{r}|\mathbf{c})\Pr(\mathbf{c})}{\sum\limits_{\mathbf{c}_j \in \mathcal{C}} p\big(\mathbf{r}\big|\mathbf{c}_j\big)\Pr\big(\mathbf{c}_j\big)} \ . \tag{2.48}$$

Therefore, in order to compute $\Pr(\mathbf{c}|\mathbf{r})$ the *a priori probabilities* of the codewords $\Pr\big(\mathbf{c}_j\big)$ must be known. However, in many cases, these probabilities are not known exactly at the receiver, making optimum decoding impossible. In this case, the best feasible decoding rule is to use the *maximum-likelihood* (ML) criterion that selects $\hat{\mathbf{c}} \in \mathcal{C}$ such as:

$$\Pr(\mathbf{r}|\hat{\mathbf{c}}) \geqslant \Pr\big(\mathbf{r}\big|\mathbf{c}_j\big) \quad , \forall \mathbf{c}_j \in \mathcal{C} \text{ with } j = 0, \ldots, 2^k - 1 \ . \tag{2.49}$$

We note that the denominator of (2.48) is independent of the signal that is transmitted. Hence, we conclude that a decoder based on the MAP criterion and one that is based on the ML criterion make the same decisions as long as the a priori probabilities $\Pr\big(\mathbf{c}_j\big)$ are all equal, i.e., all the transmitted codewords are equiprobable, as is usually the case.

2.3.2 *Hard-Decision Decoding Performance*

The only decoding algorithm for linear block codes presented so far was the syndrome's method. This algorithm assumes that the received word $\mathbf{r}$ belongs to the vector space where the code was defined, i.e. $\mathbf{r}$ is a n-tuple binary vector[10]. Sample outputs r_i of the matched filter in the receiver demodulator must therefore be quantized in two levels, denoted as "0" and "1", and the demodulator is said to make *hard-decisions*.

Decoding algorithms like the syndrome method that process the hard-decision received sequences are called *hard-decision decoding algorithms*. When the algebraic structure of the code is used to perform this decoding (as with the syndrome method, too), the algorithm is also called an *algebraic decoding algorithm* [66].

Let r_i denote an unquantized matched filter output sample, and $\bar{r}_i$ the hard decision taken on r_i according to

$$\bar{r}_i = \begin{cases} 0 & \text{, if } r_i < \tau \\ 1 & \text{, if } r_i \geqslant \tau \end{cases} , \tag{2.50}$$

where τ is the hard-decision threshold.

The optimum threshold τ is obtained using the MAP criterion [77]. For the case of coherent BPSK transmission over the AWGN channel, the MAP decision rule compares the weighted densities $p(r_i|s_i)\Pr(s_i)$ for $s_i = \pm\sqrt{E_s}$, where $p(r_i|s_i)$ is given by (2.43). Clearly, from figure 5 we conclude that the optimum threshold τ is the point at which

$$p\left(r_i \,\middle|\, s_i = \sqrt{E_s}\right) \Pr\left(s_i = \sqrt{E_s}\right) = p\left(r_i \,\middle|\, s_i = -\sqrt{E_s}\right) \Pr\left(s_i = -\sqrt{E_s}\right) , \tag{2.51}$$

and can thus be computed explicitly as

$$\tau = \frac{N_0}{4\sqrt{E_s}} \ln \frac{\Pr\left(s_i = -\sqrt{E_s}\right)}{\Pr\left(s_i = \sqrt{E_s}\right)} . \tag{2.52}$$

Usually $\Pr\left(s_i = \sqrt{E_s}\right) = \Pr\left(s_i = -\sqrt{E_s}\right)$, i.e. transmitted symbols are equiprobable, and so the decision threshold is at $\tau = 0$. Here the composite channel composed of the BPSK modulator, the AWGN channel, the demodulator and the hard-decision detector, turns out to be equivalent to a discrete-time BSC channel (as depicted in figure 6) with *crossover probability*

$$p = \Pr\left(r_i < 0 \,\middle|\, s_i = \sqrt{E_s}\right) = \Pr\left(r_i > 0 \,\middle|\, s_i = -\sqrt{E_s}\right) . \tag{2.53}$$

10 Remember that from section 2.3 on the codes are considered to be binary.

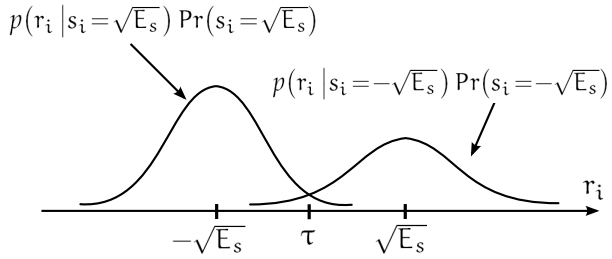


Figure 5: Optimum hard-decision threshold for BPSK transmission over the AWGN channel.

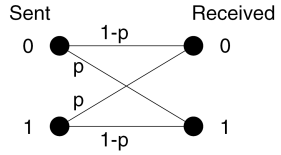


Figure 6: The binary symmetric channel model.

Hence, using equation (2.43), p is given by

$$p = \frac{1}{\sqrt{\pi N_0}} \int_0^{+\infty} e^{-\frac{(r_i + \sqrt{E_s})^2}{N_0}} dr_i = Q\left(\sqrt{\frac{2E_s}{N_0}}\right) , \tag{2.54}$$

where $Q(x)$ is the *Gaussian Q-function* that returns the area under the tail of the normalized Gaussian[11] pdf and is defined as:

$$Q(x) = \frac{1}{\sqrt{2\pi}} \int_x^{+\infty} e^{-\frac{t^2}{2}} dt . \tag{2.55}$$

The crossover probability p can be expressed as a function of the typical figure of merit E_b/N_0. Because for a binary (n, k) block code each set of n transmitted symbols ("bits") carries only k *information bits*, the average transmitted energy per information bit E_b relates to E_s according to

$$E_b = \frac{nE_s}{k} = \frac{E_s}{r_c} , \tag{2.56}$$

where r_c is the code rate defined in (2.17). Therefore, p can be expressed as

$$p = Q\left(\sqrt{2r_c \frac{E_b}{N_0}}\right) . \tag{2.57}$$

11 Gaussian distribution with zero mean and unitary standard deviation.

The hard-decision decoding metric

Considering transmission of a codeword $\mathbf{c}$ over a BSC channel, the probability of an error pattern $\mathbf{e}$ with $w_H(\mathbf{e}) = i$ occurring equals the probability that i bits of the hard-decision received word $\bar{\mathbf{r}}$ are in error and the remaining $n - i$ bits are correct. This probability denoted as $P_{\mathbf{e}}(i)$, is thus given by

$$P_{\mathbf{e}}(i) = Pr(\bar{\mathbf{r}}\,|\mathbf{c}\,) = p^i\,(1-p)^{n-i}\ . \tag{2.58}$$

Therefore, the higher the Hamming weight of the error pattern, i.e. the longer the distance $d_H(\bar{\mathbf{r}}, \mathbf{c})$, the lower the probability that it will occur[12].

As stated in section 2.3.1, an ML decoding rule selects the codeword $\hat{\mathbf{c}} \in \mathcal{C}$ that maximizes $Pr(\bar{\mathbf{r}}\,|\hat{\mathbf{c}}\,)$, which, according to (2.58), is thus equivalent to select $\hat{\mathbf{c}} \in \mathcal{C}$ that minimizes $d_H(\bar{\mathbf{r}}, \hat{\mathbf{c}})$. The Hamming distance is therefore the metric used for practical hard-decision decoding algorithms.

Probability of an undetected error

An undetected error pattern occurs whenever the received word $\bar{\mathbf{r}}$, is a codeword different from the transmitted one $\mathbf{c}$, i.e. the error pattern $\mathbf{e} = \bar{\mathbf{r}} + \mathbf{c}$ is a non-zero codeword, since the code is linear. In fact, for any error pattern $\mathbf{e} \in \mathcal{C}\backslash\{\mathbf{0}\}$, the syndrome $\bar{\mathbf{r}}\,\mathbf{H}^T$ of the received word is zero according to (2.23) and (2.36), and so no error would be detected.

Using the coefficients of the weight enumerating polynomial $A(z)$, we can thus compute exactly the probability that an undetected error will occur P_u as

$$P_u = \sum_{i=d_{min}}^{n} A_i P_{\mathbf{e}}(i) = \sum_{i=d_{min}}^{n} A_i p^i\,(1-p)^{n-i}\ . \tag{2.59}$$

Unfortunately, for most codes of practical interest the weight distribution is unknown. However, since $\mathcal{C}$ is a subspace of $GF(2)^n$, the number of codewords of weight i satisfies

$$A_i \leqslant \binom{n}{i}\ , \tag{2.60}$$

and so the following upper bound on P_u is obtained

$$P_u \leqslant \sum_{i=d_{min}}^{n} \binom{n}{i} p^i\,(1-p)^{n-i}\ . \tag{2.61}$$

12 A useful BSC channel has a crossover probability $p < 0.5$.

Probability of a decoding error

According to the syndrome method a received word $\bar{\mathbf{r}}$ is properly corrected whenever the error pattern $\mathbf{e}$ occurring during transmission is a coset leader.

Let L_i denote the number of coset leaders of weight i and l the largest Hamming weight of a coset leader in the standard array. The *probability of a correct decoding* P_c is thus given by

$$P_c = \sum_{i=0}^{l} L_i \, p^i \, (1-p)^{n-i} \, . \tag{2.62}$$

and the *probability of a decoding error* P_E, usually called *block* (or *word*) *error rate* (BLER) [66], is

$$P_E = 1 - P_c = 1 - \sum_{i=0}^{l} L_i \, p^i \, (1-p)^{n-i} \, . \tag{2.63}$$

For most (n, k) linear codes of practical interest, the syndrome method is not a feasible decoding algorithm because of its long length n, and so the weight distribution of the coset leaders is not known. However, a simple upper bound on P_E can be found. As stated in section 2.2.2, a linear code with minimum distance d_{min} is able to correct all error patterns of $t_{cor} = \lfloor (d_{min} - 1) \backslash 2 \rfloor$ or fewer errors. Therefore, the probability of the decoder committing an erroneous decoding is upper bounded by

$$P_E \leqslant \sum_{i=t_{cor}+1}^{n} \binom{n}{i} p^i \, (1-p)^{n-i} \, . \tag{2.64}$$

Equality holds in equation (2.64) if the linear block code is a perfect code. In this case, as results from (2.32), the largest Hamming weight of a coset leader in the standard array is $l = t_{cor}$ and the weight distribution of the coset leaders is known and given by

$$L_i = \binom{n}{i} \quad , \text{for } i = 0, \ldots, t_{cor} \, . \tag{2.65}$$

2.3.3 *Soft-Decision Decoding Performance*

When the sample outputs of the matched filter are left unchanged or quantized in more than two levels[13], the demodulator is said to make *soft-decisions*. Decoding

13 Once more we are restricting ourselves to the binary case, $q = 2$. For a linear code defined over GF(q), the demodulator is said to make soft-decisions whenever the number of quantization levels is greater than q.

algorithms that process the soft-decision received sequences are called *soft-decision algorithms*.

Soft-decision decoding has a higher complexity than hard-decision decoding. However, since the decoder can use the additional information contained in the unquantized (or multilevel quantized) received samples to recover the transmitted codeword, soft-decision decoding provides a much better error performance, which in practical cases can be of the order of about 2dB [66, 87].

Soft-decision decoding metrics

Metrics other than the Hamming distance must be used for soft-decision decoding. The first approach to the ML decoding of a soft-decision received sequence $\mathbf{r}$ is to use the *likelihood function*, i.e. the conditional pdf $p(\mathbf{r}|\mathbf{s})$ given by equation (2.44), or any monotonic function of it. A typical function used to simplify the computations is the natural logarithm of $p(\mathbf{r}|\mathbf{s})$, i.e.

$$\ln p(\mathbf{r}|\mathbf{s}) = -\frac{n}{2}\ln(\pi N_0) - \frac{1}{N_0}\sum_{i=0}^{n-1}(r_i - s_i)^2 . \tag{2.66}$$

The ML decoder selects the codeword $\hat{\mathbf{c}} \in \mathcal{C}$, or the equivalent BSPK modulated sequence $\hat{\mathbf{s}} = 2\hat{\mathbf{c}} - 1$, that maximizes $\ln p(\mathbf{r}|\hat{\mathbf{s}})$. According to (2.66), this is equivalent to finding the codeword BPSK sequence $\hat{\mathbf{s}}$ that minimizes the *Euclidean distance*

$$d_E(\mathbf{r}, \hat{\mathbf{s}}) = \sum_{i=0}^{n-1}(r_i - \hat{s}_i)^2 . \tag{2.67}$$

This decision rule is called the *minimum distance detection* [87] and the metric used is the Euclidean distance.

An equivalent optimum decision rule based on the ML criterion can be obtained by expanding the right-hand side of (2.67) as

$$d_E(\mathbf{r}, \hat{\mathbf{s}}) = \sum_{i=0}^{n-1} r_i^2 + n - 2\sum_{i=0}^{n-1} r_i \hat{s}_i = \|\mathbf{r}\| + n - 2\mathbf{r}\cdot\hat{\mathbf{s}} , \tag{2.68}$$

where $\mathbf{r}\cdot\hat{\mathbf{s}}$ denotes the *correlation* (*inner product*) between $\mathbf{r}$ and $\hat{\mathbf{s}}$.

In computing $d_E(\mathbf{r}, \hat{\mathbf{s}})$ over all codeword BPSK sequences $\hat{\mathbf{s}}$, the term $\|\mathbf{r}\| + n$ is constant and so it may be ignored in the computation of the metrics. Hence, minimizing $\hat{\mathbf{s}}$ among all possible codewords $\hat{\mathbf{c}} \in \mathcal{C}$ is equivalent to maximizing

the correlation between the received sequence and all $\hat{\mathbf{c}} \in \mathcal{C}$. This metric $C(\mathbf{r}, \hat{\mathbf{c}})$, called the *correlation metric*, is computed as

$$C(\mathbf{r}, \hat{\mathbf{c}}) = \mathbf{r} \cdot \hat{\mathbf{s}} = \sum_{i=0}^{n-1} r_i \hat{s}_i = \sum_{i=0}^{n-1} r_i \, (2\hat{c}_i - 1) \; . \tag{2.69}$$

Probability of a decoding error

Assume that we send a codeword $\bar{\mathbf{c}} \in \mathcal{C}$ and the received sequence is $\mathbf{r}$. Optimum soft-decision ML decoding can be performed by computing the 2^k correlation metrics of $\mathbf{r}$ with $\mathcal{C}$'s codewords. A decoding error occurs if

$$\exists \mathbf{c}_j \in \mathcal{C} \backslash \{\bar{\mathbf{c}}\} : \; C\big(\mathbf{r}, \mathbf{c}_j\big) > C(\mathbf{r}, \bar{\mathbf{c}}) \; . \tag{2.70}$$

The probability of $\bar{\mathbf{c}}$ being correctly decoded can thus be computed as

$$P_c^{(\bar{c})} = \Pr\big[C(\mathbf{r}, \bar{\mathbf{c}}) > C(\mathbf{r}, \mathbf{c}_1), \; C(\mathbf{r}, \bar{\mathbf{c}}) > C(\mathbf{r}, \mathbf{c}_2), \cdots, C(\mathbf{r}, \bar{\mathbf{c}}) > C\big(\mathbf{r}, \mathbf{c}_{2^k-1}\big)\big] \\ \text{, with } \mathbf{c}_j \in \mathcal{C} \backslash \{\bar{\mathbf{c}}\} \; , \tag{2.71}$$

which means that the probability of a correct decoding over all possible transmitted codewords is

$$P_c = \sum_{\forall \bar{\mathbf{c}} \in \mathcal{C}} P_c^{(\bar{c})} \Pr(\bar{\mathbf{c}}) \; . \tag{2.72}$$

Because the code is linear and the AWGN channel is symmetric, the error probability given by (2.71) is the same for all $\bar{\mathbf{c}} \in \mathcal{C}$. It can therefore be assumed that the all-zero codeword $\bar{\mathbf{c}}_0$ is transmitted and so

$$P_c = P_c^{(\bar{c}_0)} \; . \tag{2.73}$$

However, the computation of (2.71) is not a trivial task due to the correlations among the 2^k correlation metrics. Therefore, instead of attempting to derive the exact probability P_c, or the equivalent word error probability $P_E = 1 - P_c$, we resort to the union bound. The word error probability can thus be upper bounded as follows

$$P_E \leqslant \sum_{\mathbf{c}_j \in \mathcal{C} \backslash \{\bar{\mathbf{c}}_0\}} \Pr\big[C\big(\mathbf{r}, \mathbf{c}_j\big) > C(\mathbf{r}, \bar{\mathbf{c}}_0)\big] \tag{2.74}$$

Expanding the probability that $C\big(\mathbf{r}, \mathbf{c}_j\big) > C(\mathbf{r}, \bar{\mathbf{c}}_0)$ and using equation (2.69), we get

$$\begin{aligned}\Pr\left[C(\mathbf{r},\mathbf{c}_j) > C(\mathbf{r},\bar{\mathbf{c}}_0)\right] &= \Pr\left[\mathbf{r}\cdot\left(\mathbf{c}_j-\bar{\mathbf{c}}_0\right) > 0\right] \\ &= \Pr\left[\sum_{i=0}^{n-1} r_i\left(c_{ji}-\bar{c}_{0i}\right) > 0\right] ,\end{aligned} \tag{2.75}$$

where c_{ji} represents the bit i of the codeword $\mathbf{c}_j$.

Let S_j denote the sum $\sum_{i=0}^{n-1} r_i\left(c_{ji}-\bar{c}_{0i}\right)$. Note that S_j is a sum of independent Gaussian random variables[14] $r_i \sim \mathcal{N}\left(-\sqrt{E_s}, N_0/2\right)$ and it has only $w_H(\mathbf{c}_j)$ non-null terms that correspond to the positions where $\mathbf{c}_j$ differs from $\bar{\mathbf{c}}_0$. Hence, S_j is a Gaussian random variable with mean

$$m_{S_j} = -w_H(\mathbf{c}_j)\ \sqrt{E_s} , \tag{2.76}$$

and variance

$$\sigma^2_{S_j} = w_H(\mathbf{c}_j)\ \frac{N_0}{2} , \tag{2.77}$$

and so, the probability that $C(\mathbf{r},\mathbf{c}_j) > C(\mathbf{r},\bar{\mathbf{c}}_0)$ is given by

$$\Pr\left[C(\mathbf{r},\mathbf{c}_j) > C(\mathbf{r},\bar{\mathbf{c}}_0)\right] = \Pr\left[S_j > 0\right] = Q\left(\sqrt{2w_H(\mathbf{c}_j)\,\frac{E_s}{N_0}}\right) . \tag{2.78}$$

If $A(z)$ is the weight enumerating polynomial of the linear code $\mathcal{C}$ then, by making use of equation (2.78) and the relation (2.56) between E_s and E_b, the result is that the word error probability as a function of the signal noise ratio per information bit E_b/N_0 is upper bounded by

$$P_E \leqslant \sum_{w=d_{min}}^{n} A_w Q\left(\sqrt{2\,w\,r_c\,\frac{E_b}{N_0}}\right) . \tag{2.79}$$

For large E_b/N_0 this bound is quite tight. However, it requires the knowledge of the weight distribution of the code. A particularly useful, but somewhat looser bound can be obtained by noting that (2.78) is a monotonic decreasing function of the codeword's weight $w_H(\mathbf{c}_j)$, and so

$$P_E \leqslant \sum_{w=d_{min}}^{n} A_w Q\left(\sqrt{2\,w\,r_c\,\frac{E_b}{N_0}}\right) \leqslant \left(2^k-1\right) Q\left(\sqrt{2\,d_{min}\,r_c\,\frac{E_b}{N_0}}\right) . \tag{2.80}$$

14 Remember that it was assumed that codeword $\bar{\mathbf{c}}_0$ was transmitted and so, according to (5.99), the received samples of $\mathbf{r}$ are expressed as $r_i = -\sqrt{E_s} + n_i$.

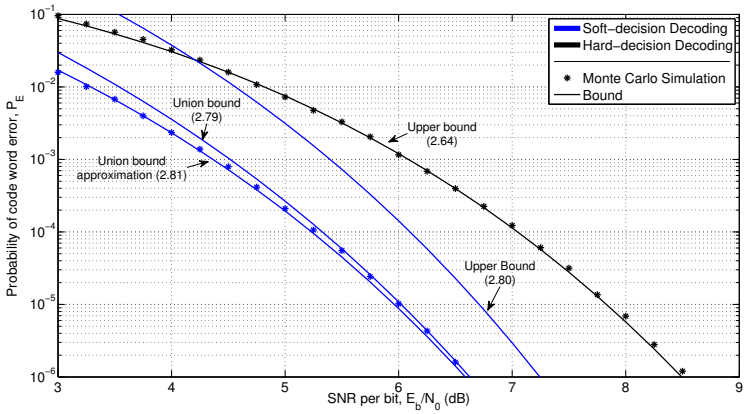


Figure 7: Comparison of bounds with exact error probability for hard-decision and soft-decision decoding for the Golay $(23, 12)$ code.

Another very good approximation of (2.79), that doesn't require the knowledge of $A(z)$, can be obtained for large values of E_b/N_0. Here the term $\sqrt{2\, d_{min}\, r_c\, E_b/N_0}$ is dominant in the summation (2.79) and so

$$P_E \leqslant \sum_{w=d_{min}}^{n} A_w Q\left(\sqrt{2\, w\, r_c\, \frac{E_b}{N_0}}\right) \approx A_{d_{min}} Q\left(\sqrt{2\, d_{min}\, r_c\, \frac{E_b}{N_0}}\right) . \qquad (2.81)$$

Bounds given by equations (2.79), (2.80) and (2.81) are compared in figure 7, with the exact word error probability (obtained by Monte Carlo simulation) for the case of the binary $(23, 12)$ Golay code described in section 2.2.4. We observe that the above mentioned properties of each of these bounds are verified. In particular, we note that (2.81) provides a quite good approximation of the exact word error probability under soft-decision decoding.

The exact performance P_E under hard-decision decoding and the numerical result obtained from the upper bound (2.64) are also plotted in figure 7. Since the Golay $(23, 12)$ code is perfect, equation (2.64) holds in equality and so gives the exact word probability error. In addition figure 7 shows that soft-decision decoding has a gain of about 2dB in performance with respect to hard-decision decoding.

2.4 LOW-DENSITY PARITY-CHECK CODES

LDPC codes also called Gallager's codes, are a special class of linear block codes whose parity-check matrix $\mathbf{H}$ has a low density of ones[15], i.e. it is *sparse*.

15 In this section, we address only the study of binary LDPC codes, and therefore we consider $\mathbf{H}$ to be a binary matrix. Non-binary LDPC codes can also be constructed, with $\mathbf{H}$ being a sparse matrix defined over $GF(q)$.

They were originally proposed in 1962 by Robert Gallager in his doctoral thesis [20, 21]. Along with turbo codes (discovered much later in 1993 by Berrou et al. [6]), LDPC codes form a family of Shannon limit capacity-approaching codes.

Unfortunately, Gallager's remarkable discovery was long before its time mainly because the *iterative message-passing a posteriori probability* (APP) decoding algorithm that he proposed was too complex for the technology of that time. As a result, LDPC codes remained largely unstudied for over 30 years, with only scattered references to them. One notable exception is the important work published in 1981 by Michael Tanner [106] in which he generalized LDPC codes and introduced a graphical representation of them, now called *Tanner graphs*. It was only in 1995 with the advent of the turbo code "revolution", that Mackay and Neal [68] rediscovered them and showed empirically [69] that near Shannon limit performance could be obtained with long LDPC-type codes and iterative decoding.

Since Mackay and Neal's rediscovery of them LDPC codes have become one of the hottest topics in coding theory. There were several reasons for the astonishing interest in Gallager's original work. In some cases LDPC codes perform even better than turbo codes [90], with their *error floor* occurring at a much lower BER. LDPC's iterative decoding algorithms are also easy to implement (with a much lower per-iteration complexity than that of turbo decoders) and are also parallelizable in hardware. In addition, the decoder declares a decoding failure in a very natural way when it is unable to correctly decode, whereas turbo decoders must perform extra computations for a stopping criterion. On the commercial side, a strong reason for the interest on them was that LDPC codes are not patent protected. But among all the advantages they have a disadvantage with respect to turbo codes: LDPC codes have a significantly higher encoding complexity since it is generically quadratic in the code dimension. This disadvantage can however be overcome by using a special class of LDPC codes named *Irregular Repeat-Accumulate* (IRA) [58], or cyclic or quasi-cyclic LDPC codes constructed based on finite geometries [62, 66].

The powerful capabilities of LDPC codes have led them to compete successfully with turbo codes and they have recently been included in several standards, such as WiMAX IEEE 802.16e [49], *Mobile Broadband Wireless Access* (MBWA) IEEE 802.20 [50], *Ethernet* IEEE 802.3an [51] and the satellite digital video broadcast DVB-S2 [19].

We use the LDPC codes defined in the WiMAX IEEE 802.16e standard to evaluate performance of the techniques of peak power control devised for higher order modulated band-limited signals, considering transmission over the AWGN channel. For this reason the special type of IRA-LDPC adopted in WiMAX is detailed in appendix B. The same is done for the LDPC codes defined in the DVB-S2 standard. Along with the PAPR study, some research was carried out on the side to study LDPC's efficient encoding and decoding hardware architectures for DVB-S2. The results of this research [30, 31, 33] are also presented in appendix B.

2.4.1 *Description of LDPC Codes: Tanner Graphs*

The original LDPC codes as proposed by Gallager [20] were *regular* LDPC codes. Gallager defined an (n, k, w_c) LDPC code as a linear block code with dimensions (n, k) characterized by a sparse parity-check matrix $\mathbf{H}$ whose columns have a uniform Hamming weight w_c and whose rows have a uniform Hamming weight $w_r = w_c \times \frac{n}{n-k}$, where $w_c \ll n-k$ (equivalent to $w_r \ll n$). Therefore, assuming $\mathbf{H}$ is full-rank, the code rate r_c relates to w_c and w_r via

$$r_c = 1 - \frac{w_c}{w_r} \, . \tag{2.82}$$

Gallager stated that by making $w_c \geqslant 3$ and following some simple rules on the design of a code parity check matrix (such as guaranteeing that no two columns of $\mathbf{H}$ have an overlap greater than one) we obtain with high probability an LDPC code with good error detecting and error correcting capabilities. In fact, Gallager proved that with $w_c \geqslant 3$ the typical minimum distance d_{min} of the ensemble of (n, k) codes that can be obtained grows linearly fast with the block size n of the code, while for $w_c = 2$, d_{min} grows only logarithmically with n [20, 80].

One may also define an *irregular* LDPC code where the Hamming weights of the columns and rows of $\mathbf{H}$ are chosen in accordance with some non-uniform distribution. Such irregular codes often outperform regular codes with similar dimensions [90]. However, this "irregularity" poses problems for the implementation in hardware, making the parallelization of the algorithms difficult.

Tanner Graphs

It is easiest to understand the sense in which an LDPC code is regular or irregular through its graphical representation. Any linear block code defined by a parity check matrix $\mathbf{H}$, can be represented by a *bipartite graph* associated with $\mathbf{H}$, known as *Tanner graph* [106].

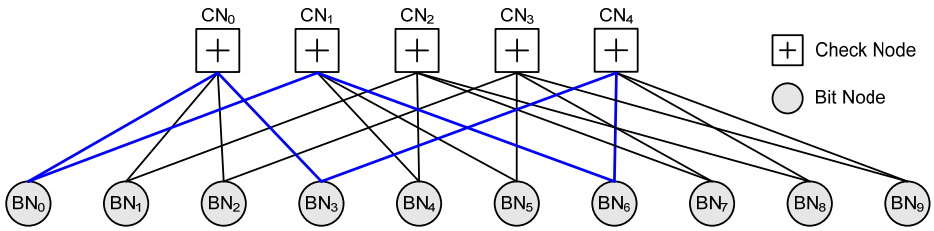
A bipartite graph is a graph [66, 89], whose *nodes* (or *vertices*) may be separated into two types, and *edges* may only connect two nodes of different types. In a Tanner graph these two types of nodes are called *variable nodes* (or *bit nodes* - BN, when the code is binary[16]) and *check nodes* (CN).

The variable and check nodes owe their name to the one-to-one correspondence between the Tanner graph and the parity check matrix $\mathbf{H}$. As we know, the $\mathbf{H}$ matrix of an (n, k) linear block code specifies a system of $n - k$ homogeneous linear equations (*check equations*) with n unknowns presented in (2.23), whose solutions are the 2^k codewords of the code. We can thus construct a graph with $n - k$ check nodes, one per each check equation, n variable nodes, one per

16 For a non-binary code some authors also call the variable nodes *symbol nodes*.

$$\mathbf{H} = \begin{array}{c} \begin{matrix} \mathrm{BN}_0 & \mathrm{BN}_1 & \mathrm{BN}_2 & \mathrm{BN}_3 & \mathrm{BN}_4 & \mathrm{BN}_5 & \mathrm{BN}_6 & \mathrm{BN}_7 & \mathrm{BN}_8 & \mathrm{BN}_9 \end{matrix} \\ \begin{bmatrix} 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 1 & 1 \end{bmatrix} \begin{matrix} \mathrm{CN}_0 \\ \mathrm{CN}_1 \\ \mathrm{CN}_2 \\ \mathrm{CN}_3 \\ \mathrm{CN}_4 \end{matrix} \end{array}$$

(a)



(b)

Figure 8: (a) The parity-check matrix $\mathbf{H}$ of a $(10, 5, 2)$ LDPC code. (b) Corresponding Tanner graph. A length-6 cycle of the Tanner graph is highlighted in blue.

each code bit[17] (symbol) c_i of a codeword $\mathbf{c}$, and where check node j, CN_j, is connected to variable node i, BN_i, whenever element h_{ji} in $\mathbf{H}$ is a 1. An example of a parity check matrix of a $(10, 5, 2)$ LDPC code and its corresponding Tanner graph is presented in figure 8.

In figure 8 a length-6 *cycle* of the Tanner graph is also highlighted. A cycle (or *loop*) of length l in a Tanner graph is a path comprising l edges and l nodes that forms a close circuit passing from one node to another without using the same edge more than once. The length of the shortest cycle existing in the graph is called the *girth*.

Cycles, especially the short ones, lead to inefficient decoding and prevent the iterative message-passing decoding algorithms, such as the *sum-product algorithm* (SPA) proposed by Gallager [20], from converging to the optimal decoding result. It is known that the sum-product algorithm yields optimum decoding when applied to cycle-free Tanner graphs [89]. But cycle-free Tanner graphs do not support good codes, with their minimum distance being $d_{min} = 2$ at code rates $r_c > 1/2$ [80]. Tanner showed that for graphs with cycles d_{min} grows with the girth of the code [106] and therefore, LDPC codes with a large girth are to be preferred. It is therefore necessary to establish some rules for the construction of LDPC codes with good distance properties and a high girth. One of the simplest rules is to guarantee that no two columns of $\mathbf{H}$ have more

17 The Tanner graph description given here is also valid for non-binary LDPC codes, where $\mathbf{H}$ is a matrix defined over $\mathrm{GF}(q)$.

than one bit in common. That way we avoid the presence of length-4 cycles (the shortest possible length of a cycle) in the Tanner graph.

There are, however, other criteria that may dictate the design approach to constructing an LDPC code, such as: structured parity check matrices for efficient encoding and decoding, near capacity performance, or low-error rate floors. Various LDPC code construction methods that target different design criteria can be found in literature [66, 80, 91]. However, this subject is beyond the scope of this thesis and so it is not addressed here.

2.4.2 *Decoding LDPC Codes: Iterative Decoding Algorithms*

An LDPC code can be decoded in various ways ranging from low to high decoding complexity and from reasonably good error performance to very good error performance. Nevertheless, almost all decoding algorithms (both hard-decision and soft-decision type algorithms) share the property of being *iterative*. LDPC hard-decision iterative decoding algorithms are usually known as *bit-flipping* algorithms [21, 66], while soft-decision algorithms used for LDPC decoding are referred to as *message-passing* algorithms [66, 77, 89, 91]. Some algorithms lie midway between hard-decision and soft-decision decoding, such as the *weight bit-flipping* [62, 66] where the simple hard-decision bit-flipping decoding is improved by including some kind of reliability information (or measure) of the received symbols in their decoding decisions. These algorithms offer a good trade-off between error performance and decoding complexity.

Besides discovering LDPC codes, Gallager also proposed both a bit-flipping and a message-passing algorithm [20, 21]. The latter, then referred to as *probabilistic decoding*, is nowadays known as *sum-product algorithm* (SPA) and is very important because it is used in areas as diverse as error control coding, signal processing and artificial intelligence[18].

The SPA algorithm forms the basis of the best known high performance LDPC decoding algorithms. As noted, the SPA is optimum when applied to cycle-free Tanner graphs [89], but also provides near optimal decoding when applied to high-girth Tanner graphs (i.e. without short length cycles, specially of length-4 and length-6), and yet it is practical to implement.

Descriptions of the SPA algorithm for decoding binary LDPC codes in the literature traditionally assume the system model defined in section 2.3, i.e. that coherent BPSK transmission is used over an AWGN memoryless channel with white spectral noise density $N_0/2$ [W/Hz], and optimum reception is achieved by matched filtering in the receiver demodulator. Therefore, in presenting the SPA algorithm, we consider the framework assumed in section 2.3 and we use the same notation, i.e. $\mathbf{c}$ denotes a codeword of the LDPC code, $\mathbf{s}$ is the

18 The SPA was discovered independently by a few researchers in an attempt to solve problems that were not always related to the decoding of channel codes. It is therefore known by different names. For example, in the field of artificial intelligence it is usually called as *belief propagation* algorithm, since it is applied in Bayesian networks.

correspondent BPSK modulated sequence, and $\mathbf{r} = \mathbf{s} + \mathbf{n}$ is the received sequence with $\mathbf{n}$ being the noise vector occurring during transmission.

Sum-product algorithm

The SPA is an iterative message-passing algorithm that operates over the code's Tanner graph. It processes the received symbols $\mathbf{r}$ iteratively to improve the reliability of each decoded code symbol c_i based on the set of restrictions (check equations) imposed by the sparse $\mathbf{H}$ matrix of the LDPC code. The main objective is to provide, at the end of each iteration, an improved estimate of the *a posteriori probability* (APP) of each codeword bit, measured as

$$\Pr(c_i = b \,|\mathbf{r}, S_i) \ , \tag{2.83}$$

where $b \in \{0, 1\}$ and S_i denotes the event that the check equations involving codeword bit c_i (i.e. BN_i) are satisfied. Hard decisions are made based on (2.83) and (2.23) is used to check if decoding was successful. If there is a decoding failure a new iteration is performed with the input being the reliability measures of code symbols computed in the previous iteration.

Because the SPA algorithm iterates over the code's Tanner graph, each node of the graph can be imagined as a processor, and the edges as message paths through which BN processors and CN processors (corresponding to the two types of nodes that form the Tanner graph) communicate. It is thus convenient to introduce the following notation:

- $J(i) = \{\text{check nodes connected to } \text{BN}_i\}$
  $= \{j : h_{ji} = 1\}$;
- $J(i)\backslash j = \{\text{ check nodes connected to } \text{BN}_i \text{ excluding } \text{CN}_j \}$;
- $I(j) = \{\text{bit nodes connected to } \text{CN}_j\}$
  $= \{i : h_{ji} = 1\}$;
- $I(j)\backslash i = \{\text{ bit nodes connected to } \text{CN}_j \text{ excluding } \text{BN}_i\}$;
- $M_{\text{CN}}(\sim j) = \{\text{ messages from all check nodes except } \text{CN}_j \}$;
- $M_{\text{BN}}(\sim i) = \{\text{ messages from all bit nodes except } \text{BN}_i \}$;
- $q_{i \to j}(b)$ = message that BN_i sends to CN_j regarding the probability that bit $c_i = b$, with $b \in \{0, 1\}$;
- $r_{j \to i}(b)$ = message that CN_j sends to BN_i regarding the probability that check node restriction j is satisfied given that $c_i = b$;
- $P_i = \Pr(c_i = 1 \,|r_i) \ .$

A BN processor receives information from the channel, and messages from its CN *neighbours*[19] as depicted in figure 9a. A BN_i processes all the received information and sends back to each of its CN neighbours (to CN_j, for instance) a message regarding its probability based on all the information it has received, excluding the information received from CN_j, i.e. only *extrinsic information* is passed to CN_j[20]. The message that BN_i passes to CN_j is therefore

$$q_{i\to j}(b) = \Pr(c_i = b \,|r_i, S_i, M_{\text{CN}}(\sim j)) \ . \tag{2.84}$$

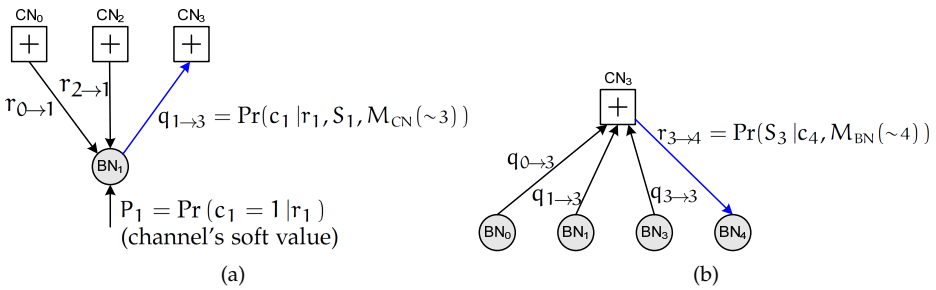


Figure 9: Subgraphs of a Tanner graph corresponding to an H matrix whose: (a) first column is $[1\,0\,1\,1\,0\cdots0]^{\mathsf{T}}_{1\times(n-k)}$; (b) third row is $[1\,1\,0\,1\,1\,0\cdots0]_{1\times n}$.

Messages received by BN_i are assumed to be statistically independent. Hence, from applying the Bayes' rule to (2.84) and making use of the independence assumption, it happens that $q_{i\to j}$ can be computed as [68, 91]:

$$q_{i\to j}(0) = k_{ij}\,(1-P_i) \prod_{j'\in J(i)\setminus j} r_{j'\to i}(0) \ , \tag{2.85}$$

$$q_{i\to j}(1) = k_{ij} P_i \prod_{j'\in J(i)\setminus j} r_{j'\to i}(1) \ , \tag{2.86}$$

where the constants k_{ij} are chosen to ensure that $q_{i\to j}(0) + q_{i\to j}(1) = 1$.

A CN processor receives only information from its BN neighbours as depicted in figure 9b. It processes the received messages in a similar manner and passes the resulting output messages to its BN neighbours with the message sent to each BN containing only information extrinsic to that BN. The message passed to each BN concerns the probability of the check node restriction being satisfied given the value of the BN and all extrinsic information, and so can be expressed as

$$r_{j\to i}(b) = \Pr\big(S_j \,|c_i = b, M_{\text{BN}}(\sim i)\big) \ . \tag{2.87}$$

19 Two nodes in a graph are said to be *neighbours* if they are connected by an edge.

20 Considering the example of figure 9a, the information that BN_0 passes to CN_2 is all the information available to BN_0 from the channel and through its neighbours, excluding the one received from CN_2.

In order for a parity-check equation involving BN_i to be satisfied, given that $c_i = 0$ (1), the remaining bits involved in the parity-check equation must contain an even (odd) number of 1s. In addition, it can be proved [91] that the probability of the sum (modulo 2, $\oplus$) of a set of independent binary digits a_i being equal to zero is given by

$$\Pr\left(\bigoplus_i a_i = 0\right) = \frac{1}{2} + \frac{1}{2}\prod_i \left[1 - 2\Pr(a_i = 1)\right] . \tag{2.88}$$

Hence, using (2.88) and assuming that messages received by CN_j are statistically independent, it follows that

$$r_{j\to i}(0) = \frac{1}{2} + \frac{1}{2}\prod_{i' \in I(j)\setminus i} \left(1 - 2q_{i'\to j}(1)\right) , \tag{2.89}$$

$$r_{j\to i}(1) = 1 - r_{j\to i}(0) . \tag{2.90}$$

Each SPA decoding iteration is thus composed of two parts following a *flooding schedule* approach [112]. A first half-iteration (*horizontal* pass) where all CN are processed and a second half-iteration (*vertical* pass) where BN are processed[21]. The algorithm is initialized by setting $q_{i\to j}(b) = \Pr(c_i = b\,|r_i)$, with $b \in \{0, 1\}$, which for the case of coherent BPSK transmission on the AWGN channel is given by

$$q_{i\to j}(b) = \frac{1}{1 + e^{\frac{r_i\,(1-2b)}{N_0}}} . \tag{2.91}$$

At the end of each iteration the pseudo a posteriori probabilities of each codeword bit are computed

$$Q_i(b) = \Pr(c_i = b\,|r_i, S_i) \tag{2.92}$$

which, using the Bayes rule and the independence assumption, is

$$Q_i(0) = k_i\,(1 - P_i) \prod_{j' \in J(i)} r_{j'\to i}(0) , \tag{2.93}$$

$$Q_i(1) = k_i P_i \prod_{j' \in J(i)} r_{j'\to i}(1) , \tag{2.94}$$

where the constants k_i are chosen to ensure that $Q_i(0) + Q_i(1) = 1$.

21 The terms *horizontal* pass and *vertical* pass are associated with the fact that we are processing all the lines and all the columns of $\mathbf{H}$.

The assumption that the messages passed to each node throughout the decoding process are statistically independent holds true only if the Tanner graph possesses no cycles. In this case, the SPA yields the exact APP [89]. However, for a graph of girth l, the independence assumption is only true up to the $l/2$-th iteration, after which messages start to loop back on themselves in the graph's cycles. Simulations have nonetheless shown that the message-passing algorithm is generally very effective provided length-4 cycles are avoided.

Algorithm 2.1 presents a summary of the SPA just described. Because the messages exchanged between a Tanner graph's nodes are probabilities this implementation is usually called the *probabilistic-domain* version of SPA.

Log-domain SPA decoder

The high number of multiplications involved in decoding an LDPC code when using the SPA probabilistic-domain approach may pose problems to an implementation in both hardware or software. Not only are additions less costly to implement than multiplications, but also a high number of multiplications of probability values can be numerically unstable. Therefore, a log-domain implementation of the SPA is to be preferred [12, 20, 47].

The SPA algorithm can be modified straightforwardly by using the concept of *log-likelihood ratio* (LLR) of a binary random variable a defined as

$$\mathrm{LLR}(a) = \ln \frac{\Pr(a=0)}{\Pr(a=1)} \,. \tag{2.95}$$

We can thus define the LLR of the a posteriori probabilities P_i, of the messages exchanged between nodes $r_{j\to i}$ and $q_{i\to j}$, and of the pseudo-probabilities a posteriori Q_i. Let:

$$\mathrm{LP}_i = \ln \frac{1-P_i}{P_i} \,, \tag{2.96}$$

$$\mathrm{Lr}_{j\to i} = \ln \frac{r_{j\to i}(0)}{r_{j\to i}(1)} \,, \tag{2.97}$$

$$\mathrm{Lq}_{i\to j} = \ln \frac{q_{i\to j}(0)}{q_{i\to j}(1)} \,, \tag{2.98}$$

$$\mathrm{LQ}_i = \ln \frac{Q_i(0)}{Q_i(1)} \,. \tag{2.99}$$

Replacing (2.85)–(2.86) in (2.98), and (2.93)-(2.94) in (2.99), after simple manipulation, we get

Algorithm 2.1 Sum-Product Algorithm

Perform each of following steps for all $(\mathrm{CN}_i, \mathrm{CN}_j)$ *pairs, i.e. for all* (i,j) *for which the code's parity check matrix* $\mathbf{H}$ *has* $h_{ji} = 1$.

STEP 0: **Initialization**

$$q_{i\to j}(0) = (1 - P_i) \ , \qquad q_{i\to j}(1) = P_i \ .$$

do {

STEP 1: **CN processing** ***(horizontal pass)***

– Update all $r_{j\to i}$ *messages:*

$$r_{j\to i}(0) = \frac{1}{2} + \frac{1}{2} \prod_{i' \in J(j)\setminus i} \left(1 - 2q_{i'\to j}(1)\right) , \qquad r_{j\to i}(1) = 1 - r_{j\to i}(0) \ .$$

STEP 2: **BN processing** ***(vertical pass)***

(a) Update all $q_{i\to j}$ *messages*

– Compute:

$$q_{i\to j}(0) = (1 - P_i) \prod_{j' \in J(i)\setminus j} r_{j'\to i}(0) \ , \qquad q_{i\to j}(1) = P_i \prod_{j' \in J(i)\setminus j} r_{j'\to i}(1) \ .$$

– Normalize computed values with $k_{ij} = \frac{1}{q_{i\to j}(0) + q_{i\to j}(1)}$:

$$q_{i\to j}(0) = k_{ij}\, q_{i\to j}(0) \ , \qquad q_{i\to j}(1) = k_{ij}\, q_{i\to j}(1) \ .$$

(b) Update the pseudo a-posteriori probabilities Q_i

– Compute:

$$Q_i(0) = (1 - P_i) \prod_{j' \in J(i)} r_{j'\to i}(0) \ , \qquad Q_i(1) = P_i \prod_{j' \in J(i)} r_{j'\to i}(1) \ ,$$

– Normalize computed values with $k_i = \frac{1}{q_i(0) + q_i(1)}$:

$$Q_i(0) = k_i Q_i(0) \ , \qquad Q_i(1) = k_i Q_i(1) \ .$$

STEP 3: **Hard-decision and stopping criterion test**

– Perform hard-decision decoding according to:

$$\hat{c}_i = \begin{cases} 1 & \Leftarrow \quad Q_i(1) \geqslant 0.5 \\ 0 & \Leftarrow \quad Q_i(1) < 0.5 \end{cases}$$

} while ($\hat{\mathbf{c}}\,\mathbf{H}^T \neq \mathbf{0}$) & (#iterations < max_iterations)

STEP 4: **Check if decoding was successful**

```
if (ĉ H^T = 0)
    return    ĉ;          // successful decoding
else
    return "Decoding failure";
end
```

$$\mathrm{Lq}_{i\to j} = \mathrm{LP}_i + \sum_{j'\in J(i)\setminus j} \mathrm{Lr}_{j'\to i} \ , \tag{2.100}$$

$$\mathrm{LQ}_i = \mathrm{LP}_i + \sum_{j'\in J(i)} \mathrm{Lr}_{j'\to i} \ . \tag{2.101}$$

Note that when the BN's processing (horizontal pass) is performed in the log-domain not only are all multiplications replaced by less costly additions, but the need for the normalization procedure in step 2 of the algorithm 2.1 disappears, and so fewer operations have to be performed.

In order to compute $\mathrm{Lr}_{j\to i}$ we can replace as well (2.89)–(2.90) in (2.97) and follow the deduction made in [21, 91]. A more interesting approach that is also easier to understand and process, makes use of the *boxplus-sum* $\boxplus$ operation proposed in [47] and defined as

$$a \boxplus b = \ln \frac{e^{a+b}+1}{e^a + e^b} \ . \tag{2.102}$$

It can be proved that the LLR of the sum (mod 2, $\oplus$) of two binary random variables (x_1 and x_2, for instance) is the boxplus-sum of their LLRs, i.e.

$$\mathrm{LLR}(x_1 \oplus x_2) = \mathrm{LLR}(x_1) \boxplus \mathrm{LLR}(x_2) \ . \tag{2.103}$$

The definition of $r_{j\to i}(b)$ given in (2.87), can be written by simple reasoning in an equivalent form as

$$r_{j\to i}(b) = \Pr\left(\bigoplus_{i'\in I(j)\setminus i} c_{i'} = b \right) \ , \tag{2.104}$$

where by the recursive application of (2.103) to (2.104) we find that $\mathrm{Lr}_{j\to i}$ can be computed as follows

$$\mathrm{Lr}_{j\to i} = \boxplus_{i'\in I(j)\setminus i} \mathrm{Lq}_{i'\to j} \ . \tag{2.105}$$

Efficient computation of $\boxplus$ can be performed by approximating (2.102) by a piecewise linear function [47]. However, a slightly rougher approximation of (2.102) should be mentioned because of its simplicity and popularity. It is known as the *min-sum* approximation [12, 47, 91] and is given by

$$a \boxplus b \approx \mathrm{sgn}(a)\,\mathrm{sgn}(b)\,\min(|a|,|b|) \ . \tag{2.106}$$

Algorithm 2.2 Log-Domain Sum-Product Algorithm and Min-Sum Algorithm

Perform each of following steps for all $(\mathrm{CN}_i, \mathrm{CN}_j)$ *pairs, i.e. for all* (i, j) *for which the code's parity check matrix* $\mathbf{H}$ *has* $h_{ji} = 1$.

STEP 0: **Initialization**

$$\mathrm{Lq}_{i\to j} = \mathrm{LP}_i \ .$$

do {

STEP 1: **CN processing** ***(horizontal pass)***

– Update all $\mathrm{Lr}_{j\to i}$ *messages:*

$$\mathrm{Lr}_{j\to i} = \boxplus_{i'\in J(j)\setminus i} \mathrm{Lq}_{i'\to j} \ .$$

Box-plus operator $\boxplus$ **definition:**

– LSPA: $$a \boxplus b = \ln \frac{e^{a+b}+1}{e^a + e^b}$$

– Min-sum: $$a \boxplus b = \mathrm{sgn}(a)\,\mathrm{sgn}(b)\,\min(|a|, |b|)$$

STEP 2: **BN processing** ***(vertical pass)***

(a) Update all $\mathrm{Lq}_{i\to j}$ *messages*

$$\mathrm{Lq}_{i\to j} = \mathrm{LP}_i + \sum_{j'\in J(i)\setminus j} \mathrm{Lr}_{j'\to i} \ .$$

(b) Update the pseudo a-posteriori probabilities LQ_i

$$\mathrm{LQ}_{i\to j} = \mathrm{LP}_i + \sum_{j'\in J(i)} \mathrm{Lr}_{j'\to i} \ .$$

STEP 3: **Hard-decision and stopping criterion test**

– Perform hard-decision decoding according to:

$$\hat{c}_i = \begin{cases} 0 & \Leftarrow \quad \mathrm{LQ} \geqslant 0 \\ 1 & \Leftarrow \quad \mathrm{LQ} < 0 \end{cases}$$

} while ($\hat{\mathbf{c}}\,\mathbf{H}^T \neq \mathbf{0}$) & (#iterations < max_iterations)

STEP 4: **Check if decoding was successful**

```
if (ĉ H^T = 0)
    return  ĉ;          // successful decoding
else
    return "Decoding failure";
end
```

Min-sum approximation causes about 1dB loss in required E_b/N_0 at BER$= 10^{-6}$ with respect to SPA. However, the simplicity of the operations involved in computations allows very efficient implementation in both hardware and software.

Algorithm 2.2 presents a summary of the outline of both the SPA in log-domain and its min-sum approach.

2.5 CONCLUSION

The basic concepts of error control coding have been presented in this chapter, with special emphasis on linear block codes. Channel coding is an integral part of any modern communication system, where it plays a crucial role in improving communications reliability. In this thesis system's FEC capabilities are explored from the standpoint of offsetting increased noise sensitivity resulting from the techniques for the envelope's peak power control that are presented in the following chapters. Derived bounds, under hard and soft-decoding allow us to analytically assess the performance of those techniques through comparison with the matched ones that are deduced in chapter 5 for the case where peak power control is used.

Some attention has also been paid to the special case of LDPC codes, as one of the most powerful families of linear codes able to approach the Shannon limit, and one of the hottest research topics in coding theory since the late 90s. They are the class of codes used throughout this thesis to evaluate the BER performance of the devised techniques of peak power control. They are also the subject of some parallel research presented in annex B.

3

MAGNITUDE MODULATION

Peak power reduction and spectral efficiency are two of the major issues in practical communications system design. As made clear in chapter 1, a high PAPR puts stringent requirements on the transmitter's DAC and HPA, increasing the cost of these systems' components and badly degrading the system's power efficiency. This problem is actually emphasized when, in order to meet the currently growing demands for higher data rates (and so higher spectral efficiency), very low roll-off root-raised cosine (RRC) pulse shaping filters and high-order constellations are used, since these lead to an increase on the signal's PAPR.

As mentioned, we have focused our study on the problem of achieving low PAPR with strictly band-limited single-carrier signals, often selected as the preferred solution for communications where the transmitter's power consumption might be a concern, such as VSAT and small mobile devices. Data *Magnitude Modulation* (MM) prior to RRC filtering is one of the techniques used to decrease PAPR in SC transmissions. The MM concept was proposed in the nineties by Miller *et al.* [74, 75], who used it to develop an adaptive peak-suppression algorithm that works well for *phase shift keying* (PSK) and generalized PSK constellations. An efficient *look-up table* (LUT) implementation of the MM technique was later proposed by Tomlinson *et al.* [4, 108] and applied with success to QPSK and OQPSK constellations.

This chapter presents an improved evolution of the LUT-MM technique. We show that is possible to use the LUT-MM concept for higher-order non-constant amplitude constellations, even considering the huge number of symbol combinations. Because RRC is driven by a sequence of data symbols varying not only in phase but also in magnitude, the control of envelope's peak-power at RRC output is more complex. In order to cope with the additional complexity, new clipping limits are proposed and the method's ability to avoid phase modulation is improved. We show, too, that LUT's computation complexity and storage requirements can be considerably reduced by exploring the constellation's symmetry and the RRC's linear phase characteristic. This chapter ends with a performance evaluation, focusing on the PAPR, back-off and BER of the improved LUT-MM technique applied to 16-QAM and 16-APSK constellations.

The content of this chapter is based mainly on [32], [34] and [38].

3.1 THE MAGNITUDE MODULATION PRINCIPLE

The underlying principles of data magnitude modulation applied to single-carrier communications are presented in this section. The generic block diagram of a typical SC transmitter is presented in figure 1 of chapter 1. In SC systems, the main contribution to PAPR results from RRC bandwidth limitation that gives rise to undesirable envelope variations in transmitted signal $x[n]$ that drives the transmitter's DAC+HPA block.

The use of *post-filtering* peak power reduction techniques may give rise to spectral spreading, causing loss of bandwidth efficiency and undesirable *adjacent channel interference*. But by making signal adjustments prior to filtering we do not affect the bandwidth of signal $x[n]$ and so guarantee operation within government allocated limits. These adjustments must, however, take into account the RRC impulse response $h[n]$ in order to prevent peak regrowth that may be caused by the filtering.

Pre-filtering PAPR reduction techniques

Regarding the latter strategy, two approaches are usually followed: *trellis coding* (TC) [13, 104], and magnitude modulation [4, 74, 75, 108]. TC techniques try to avoid critical sequences of modulated symbols that cause peak envelope excursions at the RRC's output. However, as stated before, such solutions are complex with a high computational overhead and difficult to include in currently operating radio transmission systems. In addition, they require constellation redundancy, with some bits being used for shaping control[1], which implies a reduction in the transmitted information rate.

Compared with state-of-the-art trellis PAPR reduction techniques, MM shows similar PAPR reduction gains, but offers major advantages [35]. MM operates on the modulator's output symbols and provides adjusted symbols to the RRC block, i.e. it is a simple processing block fitted between the modulator and the RRC filter. Therefore, it can be easily included in currently operating transmitter systems without replacing the equipment already installed. This is a very important advantage in terms of both system and cost. In addition, as this chapter and the next one make clear, it is possible to take advantage of the systems' existing forward-error-control (FEC) capabilities in order to efficiently compensate for the increased sensitivity to noise resulting from MM. So no further FEC is needed, which means that MM does not imply a penalty reduction in the information rate.

1 The 8-PSK trellis coding technique proposed in [104] requires the use of 1 bit per transmitted symbol, only for shaping control, in order to achieve the best performance. The maximum information rate of 8-PSK signaling is thus in fact reduced by $2/3$, corresponding in practice to the information rate of a 4-ary constellation.

3.1.1 *Magnitude Modulation*

Figure 10 illustrates the functional diagram and the basic building blocks of a typical digital communication transmitter system using the MM concept [4, 74, 75, 108]. The MM block adjusts the amplitude of the in-phase (I) and quadrature (Q) components of each modulator's output symbol

$$s[n] = s^{I}[n] + js^{Q}[n] \ , \tag{3.1}$$

so as to suppress peaks in the RRC's output signal $x[n] \in \mathbb{C}$ to be transmitted.

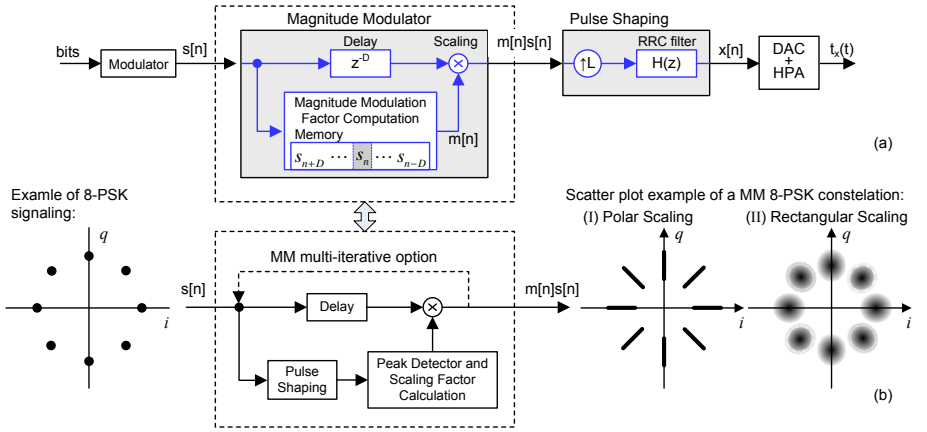


Figure 10: (a) Generic transmitter's system block diagram including MM and (b) underlying principles of MM processing.

The computation of the best MM *coefficient* $m[n]$ to apply to each symbol $s[n]$, can be accomplished by carrying out a procedure similar to the one proposed by Miller *et al.* [74, 75], that is presented in figure 10.(b). The method is composed of three steps:

1. Pulse shaping;
2. Peak detection and scaling factor calculation;
3. Scaling.

Pulse Shaping

As mentioned earlier, by performing PAPR control by signal adjustment prior to filtering, in order to prevent peak regrowth at RRC's output, the response of the pulse shaping filter to the sequence $s[n]$ must be predicted. Hence, all samples (in the neighborhood of the symbol to be magnitude modulated) that combine with the pulse shape to produce significant signal magnitude within the symbol

interval of interest must therefore be considered in the computation of the MM coefficient. That's the reason why figure 10.(a) shows a generic scheme of an MM system that looks at D past and future symbol neighbors[2] of s_n when computing the MM coefficient m_n to apply in s_n. So here a delay of DT_{SYMB} is introduced by the MM technique, with T_{SYMB} denoting the symbol duration time.

The MM's processing delay D is usually small since it depends mainly on the time duration for which the RRC impulse response maintains significant non-zero amplitude values. However, if the MM procedure is made *iterative* as shown in figure 10.(b), it increases k times, with k being the number of iterations performed per MM factor computation.

Peak detection and scaling factor calculation

However, as stated in section A.1, according to the Nyquist theory pulse shaping has to be performed at a higher sampling rate $1/T_{SMPL} = L/T_{SYMB}$, with $L \geqslant 2$ being the *up-sampler factor* of the rate expansion prior to RRC filtering. The RRC output symbol interval of interest, i.e. the one in which the contribution of symbol s_n (that is to be MM) is more prominent, it is therefore composed of L samples. This set of samples, and possibly also the ones from the neighboring intervals in its close vicinity (where the influence of s_n can still be strongly felt), are analyzed in order to detect any peak excursion exceeding a specified threshold A.

Two different types of detection may be performed: *polar* and *rectangular* analysis. Polar analysis compares the amplitude of the complex RRC output samples against the specified threshold A (*polar threshold*), while rectangular analysis monitors the amplitude of the I and Q components of those samples independently, comparing them against the *rectangular thresholds*, A_I and A_Q respectively. Rectangular thresholds are usually chosen in order to guarantee that the final MM signal $x[n]$ has an amplitude lower than the polar threshold A. However, polar analysis provides a fine control of signal magnitude.

If the threshold is exceeded, the algorithm calculates the symbol scale factor m_n to apply to s_n. A simple procedure is to scale down s_n by a factor

$$m_n = A/A_s \ , \tag{3.2}$$

where A_s is the amplitude of the maximum peak detected in the analysis interval. This value is called the *MM factor* or *MM coefficient*.

A better MM *factor* may be obtained by employing more complex algorithms, such as the one proposed by Miller *et al.* [74, 75]. For example, Miller's algorithm considers the maximum signal excursion occurring to the left and to the right

2 Hereafter, whenever there is doubt that we may be referring to a discrete time sequence $s[n]$ or to the particular sample of that sequence at instant n, we will denote the sample as s_n.

of symbol s_n during the analysis interval, and computes m_n based on the amplitude of those two peaks and their relative time distance to the symbol s_n.

Scaling

Let $h[n]$ be the impulse response of the pulse shaping filter operating at a sampling rate $1/T_{SMPL}$. Magnitude modulation is performed at symbol rate $1/T_{SYMB}$ with a scale factor m_n being applied to each symbol s_n. Scaling of the data symbols can also follow two different strategies: *polar scaling* or *rectangular scaling*.

Polar scaling multiplies both the I and Q components of symbol s_n by the same factor $m_n \in \mathbb{R}$ in order to avoid *phase modulation*. The equivalent complex digital baseband MM signal, after filtering and prior to D/A conversion, is therefore given by:

$$x[n] = \left[\sum_k m[k]\, s[k]\, \delta[n - kL] \right] * h[n] \ . \tag{3.3}$$

Alternatively, the amplitude of the I and Q components may be altered independently, which changes both the magnitude and the phase of s_n. The equivalent complex digital baseband MM signal using rectangular scaling can be expressed as

$$\begin{aligned} x[n] = &\left[\sum_k m^I[k]\, s^I[k]\, \delta[n - kL] \right] * h[n] + \\ &+ j \left[\sum_k m^Q[k]\, s^Q[k]\, \delta[n - kL] \right] * h[n] \ , \end{aligned} \tag{3.4}$$

where $m^I[k]$ and $m^Q[k]$ are the multiplier coefficients that modulate the amplitude of the transmitted pulse components $s^I[k]$ and $s^Q[k]$, respectively.

The resulting effect of magnitude modulation is the smooth scaling of the undesirable signal's peaks without clipping, as shown in figure 11 for the case of a 16-QAM constellation, with $\alpha = 0.35$ RRC filtering.

Figure 10.(b) presents a scatter plot example of an 8-PSK constellation, when using polar and rectangular scaling. As can be observed, MM can seriously increase constellation symbols' sensitivity to noise. It can even cause some symbol errors, as may happen for higher-order non-constant amplitude constellations, unless some care is taken when choosing the factor MM m_n to apply to s_n. Some simple common-sense rules are usually followed:

- Avoid phase inversion by making $m_n > 0$;
- Keep, or at least minimize the decrease in, the minimum distance between constellation symbols by choosing $m_n \in]0, 1]$.

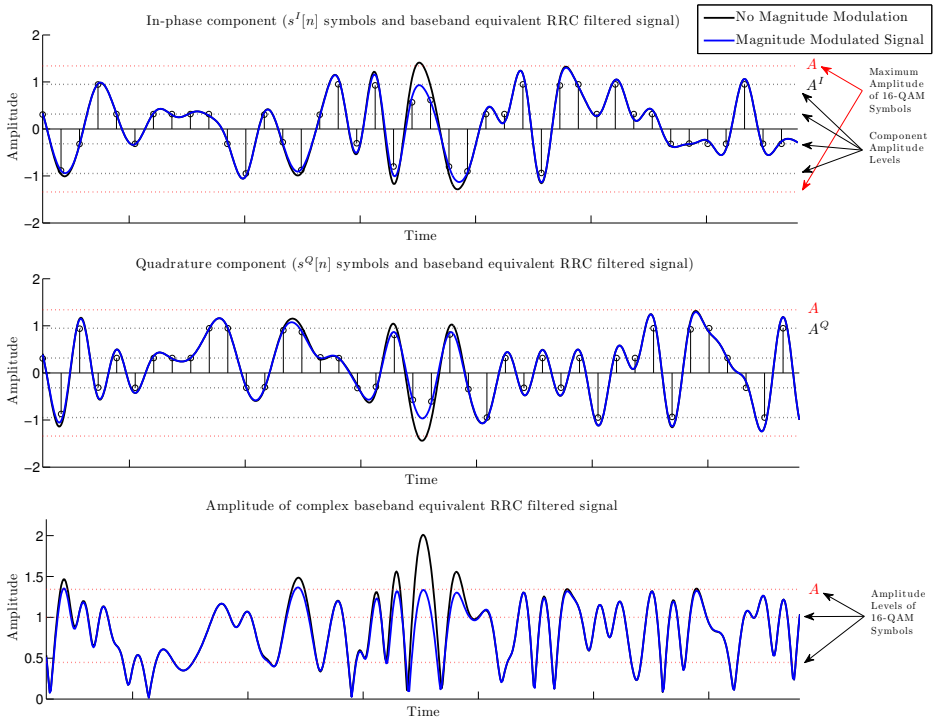


Figure 11: Sample of a $\alpha = 0.35$ RRC filtered 16-QAM signal with and without Magnitude Modulation. Rectangular scaling is used when performing MM.

Phase inversion causes symbol errors and so m_n cannot be negative. By restricting $m_n \in]0, 1]$, we only allow *attenuation* of the symbols to transmit and not their *amplification*. Amplification can be critical, especially in higher-order non-constant amplitude constellations. Here, in order to decrease some undesirable output peaks, it may be easy to amplify some *inner symbols* (small-amplitude symbols) in order to compensate for the contribution of *outer symbols* (high-amplitude symbols), while at the same time attenuating the latter and thereby minimizing the constellation's average power-loss due to MM. This minimization might even seem to be desirable based on the misleading conclusion that it would not affect the SNR at reception, and therefore system's BER performance. However, the described MM procedure would cause a dangerously close approximation of inner and outer symbols, decreasing the constellation's minimum distance too much and so increasing the constellation symbols' sensitivity to noise[3]. This would in fact seriously degrade the system's BER performance.

3 This is not the case for constant amplitude constellations, and so amplification of symbols to transmit may be allowed [74, 75].

3.2 MAGNITUDE MODULATION: LOOK-UP-TABLE BASED APPROACH

Figure 12 outlines a LUT-based implementation of the MM technique, as proposed in 2002 by Tomlinson *et al.* [4, 108]. MM coefficients are computed a priori by an iterative algorithm [32, 34], and stored in a *look-up table* (LUT).

As already noted, the optimum MM factor to apply to each data pulse $s_n = s_n^I + js_n^Q$, depends both on s_n itself, and on the D past and future symbol neighbors of s_n. Hence, MM coefficients m_n^I and m_n^Q that apply to s_n, are each stored in a LUT, at positions defined by state

$$\mathbf{S_n} = [s_{n-D} \cdots s_{n-1}\, s_n\, s_{n+1} \cdots s_{n+D}] \ . \tag{3.5}$$

When performing polar scaling MM the LUT-based magnitude modulation transmitter system present in figure 12 is simplified with only one LUT being required, since both s_n^I and s_n^Q are scaled by the same MM factor.

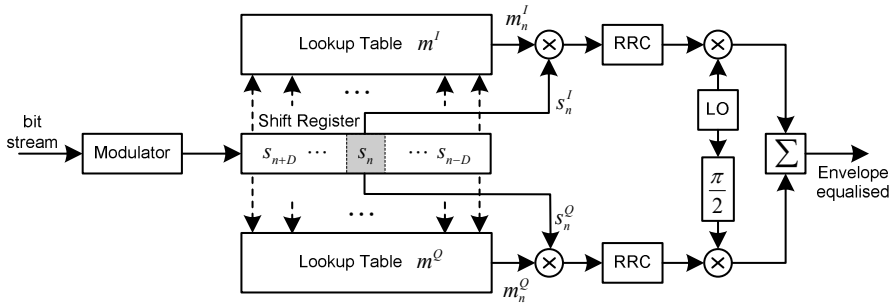


Figure 12: LUT-based magnitude modulation transmitter system.

3.2.1 *MM Look-up-Table Computation*

The iterative algorithm proposed by Tomlinson *et al.* [4, 108] for the a priori computation of MM factors simply emulates a noiseless transmission of a band-limited (RRC filtered) signal when peak-power control through clipping is used in the transmitter. The ratio of a received sample to the corresponding sent one provides the value of the MM coefficient. The method is therefore composed of the following steps presented in algorithm 3.1.

The algorithm devised by Tomlinson *et al.* was targeted to constant amplitude constellations. As mentioned, we are presenting an improved version of the LUT-MM technique, extending it to the more complex higher-order non-constant amplitude constellations.

Algorithm 3.1 Concise Algorithm for Computation of MM Factors' LUT

do {

STEP 0: **Setup input data.**

$$\textit{Input Data} = \begin{cases} \textit{Signal to be MM} & \Leftarrow \text{ first iteration} \\ \textit{MM signal from previous iteration} & \Leftarrow \text{ other iterations} \end{cases}$$

STEP 1: **Filter data using a RRC filter.**

STEP 2: **Limit the magnitude of the filtered signal to the specified threshold** A**.**

STEP 3: **Filter the resulting signal using a match RRC filter.**

STEP 4: **Sample the resulting signal to obtain the MM sequence corresponding to the input data.**

} while (Signal limitation occurs in *STEP* 2)

STEP 5: **Output the MM coefficients by performing the ratio of MM signal from the most recent iteration to the input original sequence to be MM.**

Figure 13 presents a discrete-time outline of the iterative algorithm we used for computing the MM coefficients. Different variations of this algorithm dictate the flow through the diagram. In a preferred embodiment, vectors

$$\mathbf{m^I} = \left[m^I_{-D} \cdots m^I_0 \cdots m^I_D \right] \tag{3.6}$$

and,

$$\mathbf{m^Q} = \left[m^Q_{-D} \cdots m^Q_0 \cdots m^Q_D \right] , \tag{3.7}$$

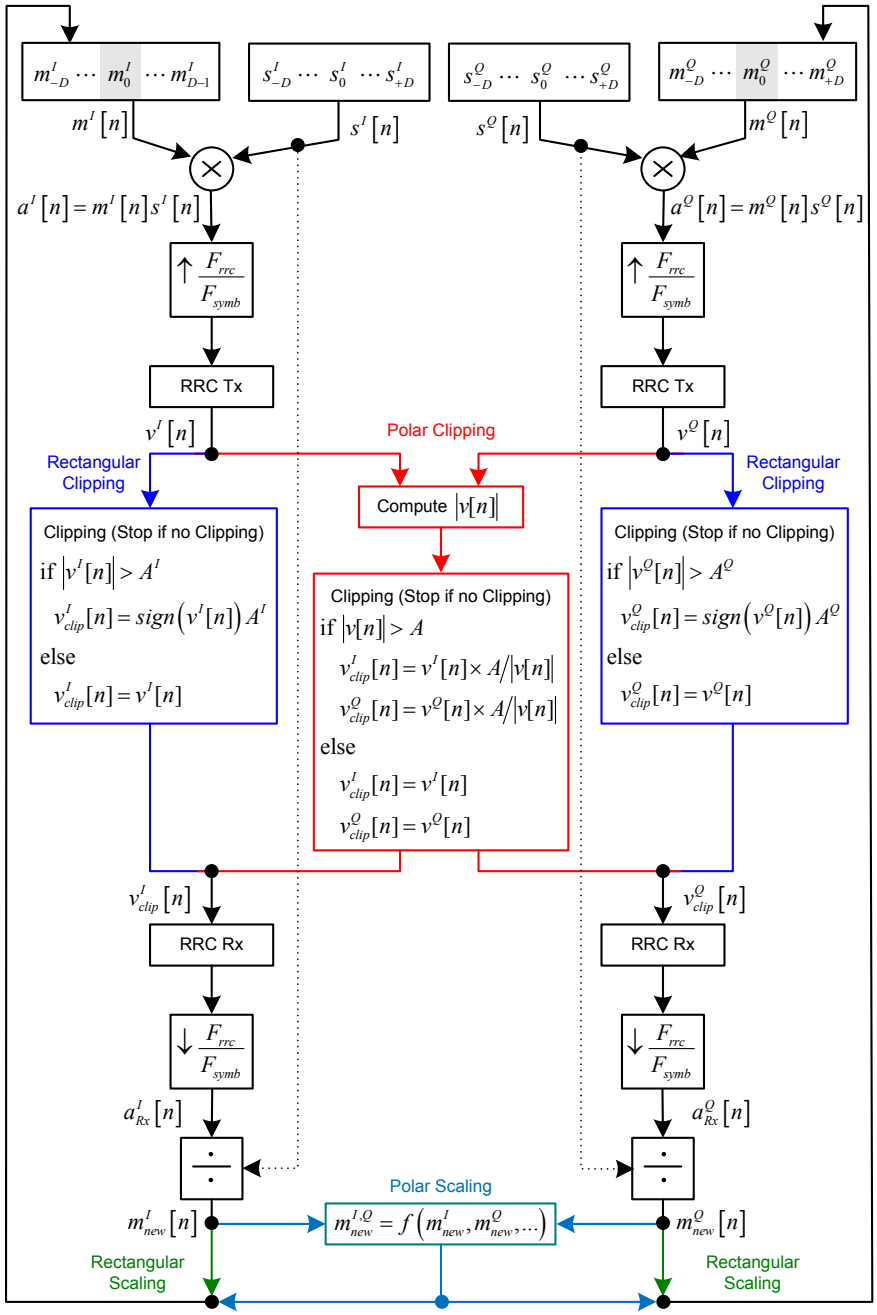
are initialized to all ones, and the algorithm iterates for each sequence of $N = 2D+1$ modulated symbols

$$\mathbf{s} = [s_{-D} \cdots s_0 \cdots s_D] , \tag{3.8}$$

until it obtains a magnitude modulated sequence

$$\begin{aligned} \mathbf{a} = & \left[m^I_{-D} s^I_{-D} \cdots m^I_0 s^I_0 \cdots m^I_D s^I_D \right] \\ & + j \left[m^Q_{-D} s^Q_{-D} \cdots m^Q_0 s^Q_0 \cdots m^Q_D s^Q_D \right] \end{aligned} \tag{3.9}$$

Figure 13


$m^I_{-D} \cdots m^I_0 \cdots m^I_{D-1}$
$s^I_{-D} \cdots s^I_0 \cdots s^I_{+D}$
$s^Q_{-D} \cdots s^Q_0 \cdots s^Q_{+D}$
$m^Q_{-D} \cdots m^Q_0 \cdots m^Q_{+D}$
$m^I[n]$
$s^I[n]$
$s^Q[n]$
$m^Q[n]$
$a^I[n] = m^I[n]s^I[n]$
$a^Q[n] = m^Q[n]s^Q[n]$
$\uparrow \frac{F_{rrc}}{F_{symb}}$
RRC Tx
$v^I[n]$
$v^Q[n]$
Polar Clipping
Rectangular Clipping
Rectangular Clipping
Compute $|v[n]|$
Clipping (Stop if no Clipping)
if $|v^I[n]| > A^I$
$v^I_{clip}[n] = sign(v^I[n])A^I$
else
$v^I_{clip}[n] = v^I[n]$
Clipping (Stop if no Clipping)
if $|v[n]| > A$
$v^I_{clip}[n] = v^I[n] \times A/|v[n]|$
$v^Q_{clip}[n] = v^Q[n] \times A/|v[n]|$
else
$v^I_{clip}[n] = v^I[n]$
$v^Q_{clip}[n] = v^Q[n]$
Clipping (Stop if no Clipping)
if $|v^Q[n]| > A^Q$
$v^Q_{clip}[n] = sign(v^Q[n])A^Q$
else
$v^Q_{clip}[n] = v^Q[n]$
$v^I_{clip}[n]$
$v^Q_{clip}[n]$
RRC Rx
RRC Rx
$\downarrow \frac{F_{rrc}}{F_{symb}}$
$a^I_{Rx}[n]$
$a^Q_{Rx}[n]$
Polar Scaling
$m^I_{new}[n]$
$m^{I,Q}_{new} = f(m^I_{new}, m^Q_{new}, \ldots)$
$m^Q_{new}[n]$
Rectangular Scaling
Rectangular Scaling

that guarantees that no limiting of the RRC filtered signal $v[n]$ (see figure 13) is necessary, i.e. $|v[n]|$ is upper bounded by

$$|v[n]| \leqslant A \ , \tag{3.10}$$

where A is the polar threshold, chosen so as to guarantee that the transmitter's HPA is not driven into saturation[4].

At the end of the iterative procedure, the m_0^I and m_0^Q values are stored in the look-up table at the position defined by state $\mathbf{s}$ (i.e. mapped in the corresponding bit sequence).

Note, however, that during the iterative procedure all neighboring symbols of s_0 are also magnitude modulated (pre-distorted), and not only s_0, i.e. all elements of the $\mathbf{m^I}$ and $\mathbf{m^Q}$ vectors are updated. This ensures that final values for m_0^I and m_0^Q will also take into consideration the magnitude modulation pre-distortion that neighboring symbols may suffer. Hence, applying the obtained MM coefficients m_0^I and m_0^Q to symbol s_0 when state $\mathbf{s}$ occurs in the transmitted stream guarantees that minimum excursions above the threshold voltage A due to s_0 arise in the corresponding RRC output symbol interval of interest and its adjacent symbol intervals.

3.2.2 *Filtering finite length signals*

Border problems for each state $\mathbf{s}$ due to RRC filtering of a finite length signal with zero padding [76], have a small impact on the accuracy of magnitude modulation values. During the iterative procedure, the MM factor of a border symbol (s_{-D} or s_D) may need to be changed in order to avoid clipping because of lack of neighboring information. That change will be propagated to all inner symbols of the state. Hopefully, final coefficients can be refined by reducing the N look-up table to a new one with $N' = N-2$ symbols per state, through a simple averaging operation.

Denote by $\mathcal{X}$ the set of constellation symbols of a M-ary constellation, $\mathbf{s}^{(N)}$ an N-tuple state defined by $N = 2D+1$ symbols as in (3.8) and let

$$\mathbf{m}_{\mathbf{s}^{(N)}} = \left(m^I_{\mathbf{s}^{(N)}}, m^Q_{\mathbf{s}^{(N)}} \right) \tag{3.11}$$

be the pair of MM coefficients resulting from applying the LUT-MM algorithm of figure 13 to state $\mathbf{s}^{(N)}$.

4 Infinite loop situations may happen, however, when peak power of $v\,[n]$ falls to the close vicinity of the polar threshold A^2, but remains higher than A^2. Therefore, in order to increase the algorithm's robustness, the limit condition (3.10) should be accompanied by monitoring $v\,[n]$ peak power of $v[n]$. The iterative procedure stops when $\max|v\,[n]|$ is in the close vicinity of A and does not decrease from one iteration to the next.

Let $N' = N-2$ and $\mathbf{s}^{(N')}$ be the state formed by the N' middle symbols of a state $\mathbf{s}^{(N)}$. There are M^2 states $\mathbf{s}_{i,j}^{(N)}$ that give rise to the same state $\mathbf{s}^{(N')}$, where

$$\mathbf{s}_{i,j}^{(N)} = \left[\; s_i \;\; \vdots \;\; \mathbf{s}^{(N')} \;\; \vdots \;\; s_j \; \right] \quad , \text{with} \;\; s_i, s_j \in \mathcal{X} \, . \tag{3.12}$$

A pair of refined MM coefficients $\mathbf{m}_{\mathbf{s}^{(N')}}$ for state $\mathbf{s}^{(N')}$ can be obtained by taking an average over all $\mathbf{m}_{\mathbf{s}_{i,j}^{(N')}}$ with $i, j = 0, \cdots, M-1$, as follows

$$\mathbf{m}_{\mathbf{s}^{(N')}} = \frac{1}{M^2} \sum_{i=0}^{M-1} \sum_{j=0}^{M-1} \mathbf{m}_{\mathbf{s}_{i,j}^{(N')}} \, . \tag{3.13}$$

With this procedure we have nearly succeeded in eliminating the boundary influence. With reference to (3.5), note that N should be large enough for the computed pair of MM factors that applies to a symbol s_n to take into account both the s_n itself and the influence of its neighbors. If this average procedure is applied recursively, although the boundary influence on the computed MM coefficients vanishes and the LUT memory size falls, the lack information from s_n neighbors causes a decrease of the MM envelope control performance. Figure 14 shows these effects. It plots the envelope maximum excursion and the corresponding required back-off (dB) of a BPSK transmission with $\alpha = 0.35$ RRC pulse-shaping, as a function of the LUT index size (bits) when: the LUT-MM algorithm from figure 13 is used alone; the LUT-MM algorithm is used followed by single-step and a two-step reduction through averaging of the LUT size; and the recursive application of the averaging procedure. Excursion values are normalized assuming BPSK symbols $s_n = \pm 1$ and $G_{RRC} = 1$. There is a clear reduction in the required back-off as a function of N (LUT index size) with the use of the MM technique after elimination of the boundary influence through (3.13).

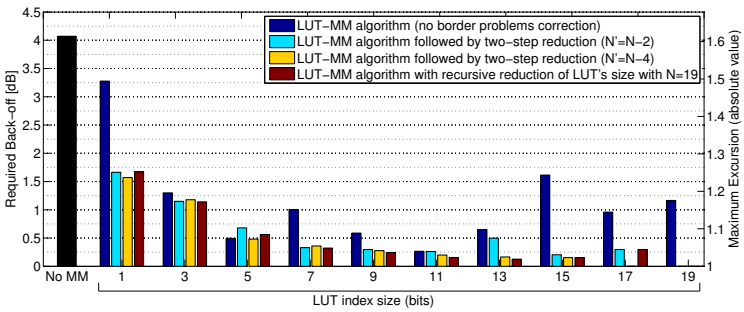


Figure 14: Envelope peak power reduction *vs.* LUT index size.

3.2.3 *LUT-MM Improvements to Handle Generic M-ary Constellations*

Figure 13 shows four algorithm approaches for the computation of magnitude modulation tables:

- Rectangular Clipping - Rectangular Scaling (RC-RS);
- Rectangular Clipping - Polar Scaling (RC-PS);
- Polar Clipping - Rectangular Scaling (PC-RS);
- Polar Clipping - Polar Scaling (PC-PS).

Clipping

According to algorithm 3.2, envelope clipping after RRC filtering at the transmitter can be performed by separately limiting the v_n^I and v_n^Q components (*rectangular clipping*) or by rescaling them together, when the envelope's magnitude $|v[n]| > A$ (*polar clipping*).

In [4], the authors chose the component clipping levels as

$$A^I = A^Q = \frac{A}{\sqrt{2}} \tag{3.14}$$

for the rectangular clipping case, since they were dealing with constant amplitude constellations.

For non-constant amplitude modulations different values must be set for A^I and A^Q according to the constellation design. Denoting the M symbols of an M-ary constellation by X_n, with $n = 0, \cdots, M-1$, we can guarantee that $|x[n]| \leqslant A$ in the separate limiting case of v_n^I and v_n^Q by setting

$$A^I = \max_{i=0,\cdots,M-1} |\Re(X_n)| \ , \tag{3.15}$$

$$A^Q = \max_{i=0,\cdots,M-1} |\Im(X_n)| \ , \tag{3.16}$$

where $\Re(X_n)$ and $\Im(X_n)$ denote the real and imaginary parts of $X_n \in \mathbb{C}$, and, thus,

$$A = \max_{i=0,\cdots,M-1} |X_n| = \sqrt{(A^I)^2 + (A^Q)^2} \ . \tag{3.17}$$

These conditions are simply an extension of the limits defined for constant amplitude modulations. However, they are very restrictive and MM coefficients computed according to them lower the power of the transmitted signal too much, leading to a decline in the BER performance. For M-ary non-constant

amplitude modulations and equiprobable symbols, the probability that one of the signal components, v_n^I or v_n^Q, violates (3.15) and (3.16) while $|x[n]| \leqslant A$, is very high, especially, if the majority of constellation symbols have amplitude lower than A. Therefore, the clipping limits should take into account the average energy of both I and Q components according to

$$A^I = \sqrt{A^2 - E\left\{|\Im(X_n)|^2\right\}}, \tag{3.18}$$

$$A^Q = \sqrt{A^2 - E\left\{|\Re(X_n)|^2\right\}}, \tag{3.19}$$

where $E\{X_n\}$ is the expected value of a random variable X_n.

These new limits, do not guarantee the condition $|x[n]| \leqslant A$ at every instant. In fact, in worst case conditions a maximum excursion of

$$\sqrt{(A^I)^2 + (A^Q)^2} > A \tag{3.20}$$

may be observed, but, their probability is too low when compared to the conditions expressed by (3.15) and (3.16).

Scaling

In order to avoid phase modulation m^I and m^Q can be forced to be equal, $m^{I,Q}$ (*polar scaling*). In [4], the authors proposed using the arithmetic mean, i.e. in figure 13, $m_{new}^{I,Q}$ was defined as

$$m_{new}^{I,Q} \triangleq \frac{m_{new}^I + m_{new}^Q}{2}. \tag{3.21}$$

for the QPSK case. For higher order constellations whose symbols have unequal amplitude components we define a new averaging function. The idea is to preserve the power of the MM symbol when restricting $m^I = m^Q$, compared with the power of the symbol obtained with rectangular scaling. Hence, $m_{new}^{I,Q}$ is computed as

$$m_{new}^{I,Q} \triangleq \sqrt{\frac{(m_{new}^I s^I)^2 + (m_{new}^Q s^Q)^2}{(s^I)^2 + (s^Q)^2}}. \tag{3.22}$$

3.2.4 *LUT-MM Complexity Discussion and Algorithm Optimization*

Computational complexity and look-up tables size (i.e. storage requirements) depend on the different clipping and scaling approaches that can be followed,

the shape and dimension of the considered M-ary constellation and the shift register length, $N = 2D + 1$ (figure 13).

Denote by b_I and b_Q the number of bits associated with I and Q components of the constellation symbols. The number of constellation symbols is

$$M = 2^{b_I + b_Q} , \tag{3.23}$$

and for square constellations,

$$b_I = b_Q = \frac{\sqrt{M}}{2} . \tag{3.24}$$

From figure 13 it is easy to conclude that for the RC-RS case the two computation branches are completely separated, with m_0^I and m_0^Q depending only on $\mathbf{s}^I = \Re(\mathbf{s})$ and $\mathbf{s}^Q = \Im(\mathbf{s})$, respectively. In the remaining cases there are interconnections between the two branches and so the m_0^I and m_0^Q values depend on the whole vector $\mathbf{s}$. These correspond to a huge difference in the LUT's size between the RC-RS approach and the remaining cases as shown in table 3.

	RC-RS CASE	REMAINING CASES
table m^I:	$2^{b_I N}$	$M^N = 2^{(b_I + b_Q)N} = 2^{b_I} \times 2^{b_Q N}$
table m^Q:	$2^{b_Q N}$	$M^N = 2^{(b_I + b_Q)N}$

Table 3: Look-up tables length for different LUT-MM algorithm approaches

If we consider a square constellation, the complexity of the RC-RS case can be simplified even more. The set of combinations $\mathbf{s}^I$ and $\mathbf{s}^Q$ to compute the look-up tables for m_0^I and m_0^Q are equal. Thus only one branch of the diagram (figure 13) needs to be run, so the magnitude modulation system (figure 12) can be implemented with a single look-up table.

Exploring constellation symmetry

The main problem posed by extending the method to a higher order M-ary constellation is the high number of constellation symbols of non-constant amplitude, which leads to a huge number of combinations to be considered, even for a system with a small memory. Hopefully, the M-ary constellations considered are usually mutually symmetric with respect to both the real and imaginary axis.

As stated before, let $\mathcal{X}$ denote the set of constellation symbols of a M-ary constellation and $\mathcal{X}^{n^{th}}$ the subset of the ones belonging to the n^{th} quadrant. For

a symmetric constellation, any state $\mathbf{s}$ defined as in (3.8), whose central symbol $s_0 \notin \mathcal{X}^{1^{st}}$, i.e.

$$\arg(s_0) = \alpha + n\frac{\pi}{2} \quad , \text{ with } \alpha \in \arg\left(\mathcal{X}^{1^{st}}\right) \text{ and } n \in \{1,2,3\} , \tag{3.25}$$

can be transformed into an equivalent state

$$\mathbf{s}' = \left[s'_{-D} \cdots s'_0 \cdots s'_D\right] = e^{-jn\frac{\pi}{2}}\mathbf{s} \tag{3.26}$$

where $s'_0 \in \mathcal{X}^{1^{st}}$ and $\arg(s'_0) = \alpha$.

Figure 15 presents an example of a 16-QAM equivalent symbol state through the rotation procedure (3.26).

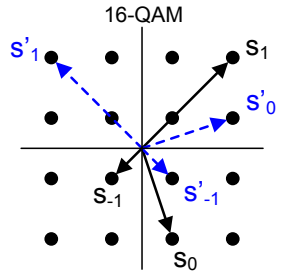


Figure 15: Equivalent $\frac{n\pi}{2}$ rotated state of $\mathbf{s} = [s_{-1}\ s_0\ s_1]$.

Based on the following proposition, only states whose central symbol belongs to $\mathcal{X}^{1^{st}}$ need to be computed and stored in a LUT. This leads to a reduction in the number of states to be considered, by a factor of 2 for RC-RS and by a factor of 4 for the remaining LUT-MM approaches.

Proposition 3.1. *For rotated states* $\mathbf{s}' = e^{-jn\frac{\pi}{2}}\mathbf{s}$ *(with* $\mathbf{s}' \in \mathcal{X}^{1^{st}}$ *and* $\mathbf{s} \notin \mathcal{X}^{1^{st}}$*), magnitude modulation factors are related by:*

$$s_0 \in \left\{\mathcal{X}^{2^{nd}},\ \mathcal{X}^{4^{th}}\right\} \quad \Rightarrow \quad \left(m^I,\ m^Q\right) = \left(m'^Q,\ m'^I\right) \tag{3.27}$$

$$s_0 \in \mathcal{X}^{3^{rd}} \quad \Rightarrow \quad \left(m^I,\ m^Q\right) = \left(m'^I,\ m'^Q\right) \tag{3.28}$$

Proof. Accordingly with (3.26):

$$\mathbf{s} = e^{jn\frac{\pi}{2}}\mathbf{s}' \Rightarrow \begin{cases} \mathbf{s}^I = -\mathbf{s}'^Q & , \mathbf{s}^Q = \mathbf{s}'^I \quad \Leftarrow \quad n = 1 \\ \mathbf{s}^I = -\mathbf{s}'^I & , \mathbf{s}^Q = -\mathbf{s}'^Q \quad \Leftarrow \quad n = 2 \\ \mathbf{s}^I = \mathbf{s}'^Q & , \mathbf{s}^Q = -\mathbf{s}'^I \quad \Leftarrow \quad n = 3 \end{cases} \tag{3.29}$$

The symmetry of the LUT-MM algorithm 3.2 diagram (figure 13) leads to the conclusion that swapping $\mathbf{s'}^{I}$ and $\mathbf{s'}^{Q}$ components at the algorithm input produces the same, but swapped, MM factors at output. Also, a sign change of vector $\mathbf{s'}^{I}$ and/or $\mathbf{s'}^{Q}$ propagates through the respective algorithm branch until it is cancelled at the division block and so, MM factors for $-\mathbf{s'}^{I}$ and $-\mathbf{s'}^{Q}$, are the same as for $\mathbf{s'}^{I}$ and $\mathbf{s'}^{Q}$, respectively.

Based on these properties, (3.27) and (3.28) results directly from (3.29). □

Exploring RRC symmetry

The RRC digital filters in figure 12 and figure 13 are typically chosen to be linear phase which allows the number of combinations to be computed according to proposition 3.2 to be reduced further.

Proposition 3.2. *Let* $\mathbf{s_r}$ *be the time reversal of state* $\mathbf{s}$*. For linear phase filters with even impulse response*

$$h_{rrc}[n] = h_{rrc}[-n] \ , \tag{3.30}$$

the two states, $\mathbf{s_r}$ *and* $\mathbf{s}$*, have the same magnitude modulation coefficients.*

Proof. Data sequences corresponding to states $\mathbf{s}$ and $\mathbf{s_r}$ are related by

$$s_r[n] = s[-n] \, . \tag{3.31}$$

According to the LUT-MM algorithm 3.2, it results

$$a_r[n] = a[-n] \, . \tag{3.32}$$

and so, considering (3.30), we conclude that

$$v_r[n] = a[-n] * h_{rrc}[n] = v[-n] \ . \tag{3.33}$$

Thus, making $\mathbf{s_r}$ the input, signals at each step of the algorithm are just time reversed versions of the corresponding signals with $\mathbf{s}$ as input, and so, LUT-MM algorithm output is the same.

□

Proposition 3.2 allows the number of states to be computed to be reduced to almost half (there are $M^{\frac{N-1}{2}}$ states for which $\mathbf{s_r} = \mathbf{s}$).

Exploring both the constellation symmetries and RRC linear phase, the LUT-MM algorithm complexity and look-up table storage requirements are reduced by almost 4 times for RC-RS, and 8 times for the remaining cases, as shown in table 4.

	RC-RS CASE	REMAINING CASES
table m^I:	$\frac{1}{4}\left(2^{b_I N} + 2^{b_I\left(\frac{N+1}{2}\right)}\right)$	$\frac{1}{8}\left(M^N + M^{\frac{N+1}{2}}\right)$
table m^Q:	$\frac{1}{4}\left(2^{b_Q N} + 2^{b_Q\left(\frac{N+1}{2}\right)}\right)$	$\frac{1}{8}\left(M^N + M^{\frac{N+1}{2}}\right)$

Table 4: Computation complexity and LUT storage requirements considering constellation and RRC symmetries.

3.2.5 *Simulation Results*

This section presents some experimental results showing that is possible to extend the MM technique to constellations with $M = 16$. A 16-APSK constellation with $\gamma = 3.15$, taken from DVB-S2 standard [19], and a 16-QAM constellation are considered. Both constellations are normalized, i.e. with average energy per transmitted symbol equal to one.

Figure 16 ilustrates a sample of an $\alpha = 0.35$ RRC filtered 16-APSK signal with and without MM. The method clearly limits the maximum peak power. Envelope excursions above maximum amplitude A, are smooth scaled without clipping, as desired. In practice, with a restriction on the maximum admissible power at HPA input to prevent HPA output saturation, the MM method allows the transmission of much higher power to the channel, as can be observed in figure 16.(a) on the right, where the outer constellation symbols are closer to the limit circle.

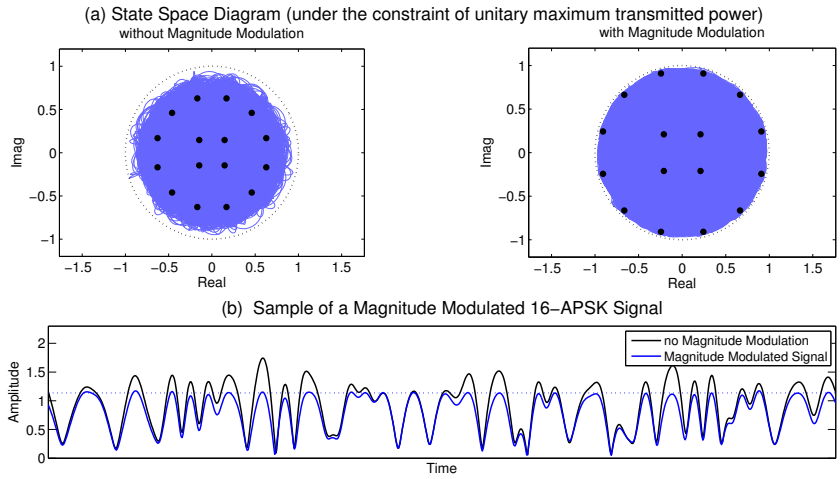


Figure 16: Sample of an $\alpha = 0.35$ RRC filtered 16-APSK signal with and without magnitude modulation.

For the polar scaling case, the new proposed function (3.22) is evaluated in comparison with the (3.21) used previously [4]. The results presented in figure 17 consider a system memory of $N = 5$ and polar clipping for the 16-APSK constellation. Performance is also evaluated in terms of BER for the AWGN channel, assuming the linear transmission (ideal HPA operation) of an $\alpha = 0.35$ RRC pulse shaped signal. Forward error correcting coding was used, considering the $(576, 288)$ LDPC code defined in the WiMAX 802.16e standard [49]. The logarithmic SPA algorithm described in section 2.4.2, following a *horizontal schedule* [29, 100], approach, was used on decoding.

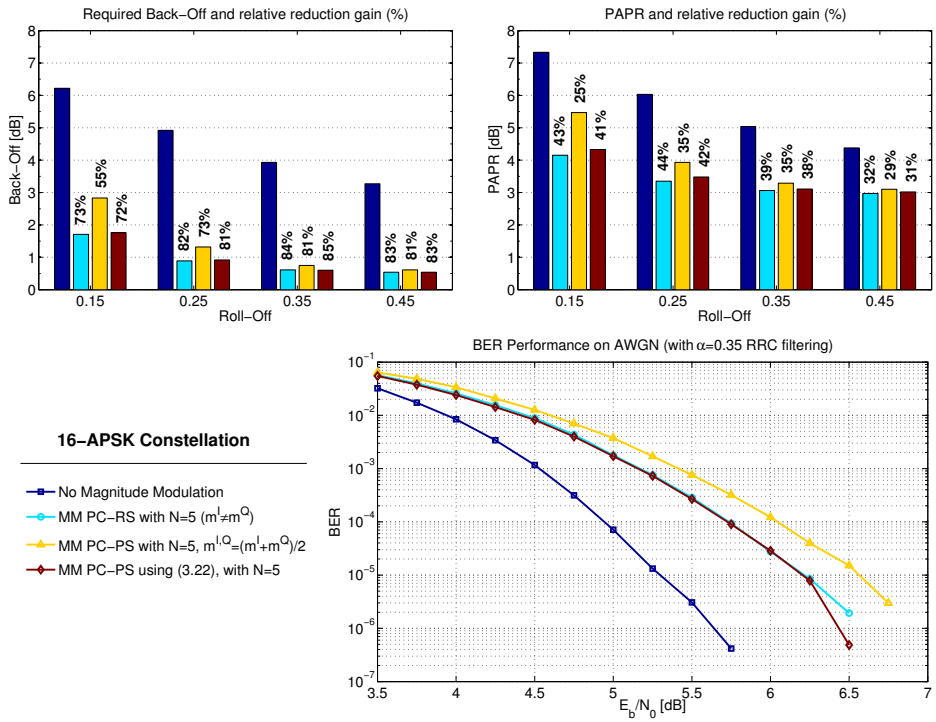


Figure 17: BER performance, PAPR and back-off gains of several LUT-MM approaches for a 16-APSK RRC filtered signal.

Even for a small memory system, e.g. $N = 5$, the LUT-MM method allows a considerable reduction of about 80% in the required back-off. PAPR reduction in dB is lower than the back-off reduction as desired, i.e. the expected decrease on transmitted average power due to the MM procedure is much lower than the desired reduction in output peak power, and so a small penalty on BER performance should be expected. This is confirmed by the BER results. No further FEC is needed. In fact, the short LDPC code efficiently offsets the increased sensitivity to noise due to the MM technique, with the MM system performing

only 1dB away (at BER= 10^{-6}) compared with the linear transmission non-MM case.

A useful measure to access the overall performance of an MM technique is the net back-off gain (dB), which is defined at a given BER per [36]

$$\text{net\_}G_{BO} = G_{BO} - L_{@BER} \quad (\text{dB}) \,, \tag{3.34}$$

where G_{BO} denotes the back-off reduction gain resulting from applying MM, and $L_{@BER}$ denotes BER performance loss measured in (dB) with respect to the linear non-MM system at a given BER.

We observe in figure 17 that net_G_{BO} is more than 2.5dB at BER= 10^{-6}, which is a significant improvement.

Considering function (3.22), it presents no loss compared with the LUT-MM method under the rectangular scaling approach, i.e. $m^I \neq m^Q$, with the major advantage of not introducing phase distortion in transmitted symbols that could cause synchronization problems on reception [72, 73]. In the typical range of interest, $\alpha \in [0.15; 0.35]$, the back-off gain is significantly better than that obtained by the average function (3.21) [4]. Figure 17 shows also that even at

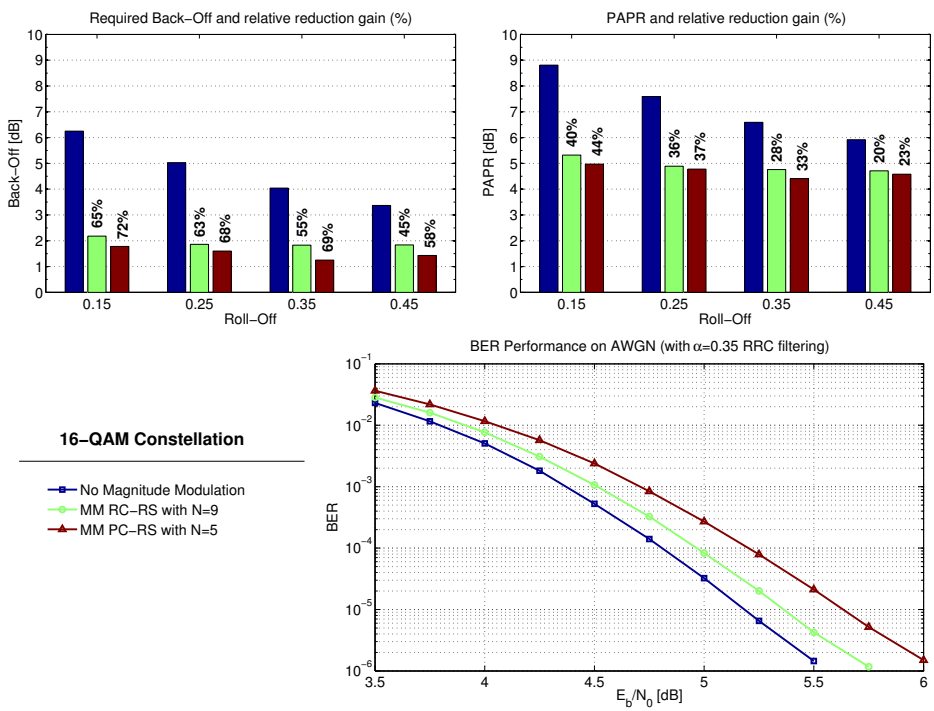


Figure 18: BER performance, PAPR and back-off gains of several LUT-MM approaches for a 16-QAM RRC filtered signal.

$\alpha = 0.35$, where both functions present similar back-off gains, a further gain of about 0.4dB is obtained at BER= 10^{-5}.

We also evaluated the advantage of using a RC-RS approach when separation can be performed on the set of bits that map to I and Q symbol components. Results for the 16-QAM modulation for the RC-RS and PC-RS approaches are presented in figure 18. Here, both methods are compared within a similar LUT size criterion. Performance considering a system memory of $N = 9$ for RC-RS is compared with PC-RS with $N = 5$.

As desired, both methods show a significant reduction of the back-off ($60\% - 70\%$ range). Although the performance for PC-RS with $N = 5$ is slightly better than for RC-RS with $N = 9$, there is no flexibility for changing N because of the problematic huge LUT size. In addition, in terms of BER RC-RS performs better than PC-RS, only 0.2dB apart from the non-MM system at BER= 10^{-6}, which also confirms the efficiency of the LUT-MM method.

BER results also show that it is possible to take advantage of the systems' existing forward-error-control (FEC) capabilities in order to efficiently compensate the increased sensitivity to noise resulting from MM. With respect to the linear transmission non-MM case a much lower BER performance degradation is observed compared with the obtained reduction in the required back-off. The net_G_{BO} is 2dB for RC-RS and 2.3dB for PC-RS at BER= 10^{-6}.

3.3 CONCLUSION

The concept of magnitude modulation MM has been presented. The LUT-MM approach, formerly proposed by Tomlinson *et al.* [4, 108] for QPSK modulation formats, has been improved and extended to constellations with dimension up to $M = 16$. In order to cope with the additional complexity, new clipping limits have been proposed and the method's ability to avoid phase modulation has been improved. A detailed analysis of the different approaches to compute the magnitude modulation tables has been performed. The constellation's symmetry and the RRC's linear phase characteristic have also been explored so as to compact the look-up tables and speed-up their computation.

Finally, the improved LUT-MM technique as applied to 16-QAM and 16-APSK constellations was evaluated, with the focus on PAPR, back-off and BER. Net back-off gains (net_G_{BO}) above 2dB were reported (at BER= 10^{-6}) when considering an $\alpha = 0.35$ roll-off RRC filter and a $(576, 288)$ LDPC code.

4

POLYPHASE MAGNITUDE MODULATION

This chapter presents a novel *Multistage Polyphase Magnitude Modulation* (MPMM) scheme for efficiently controlling the envelope's peak power of single carrier band limited signals. Building on the *polyphase decomposition* of the *finite impulse response* (FIR) of the pulse shaping filter, a new LUT-less approach of the magnitude modulation's concept is developed. Contrary to the traditional MM solutions based on LUTs, in the MPMM technique the multiplier coefficients that magnitude modulate the data stream prior to RRC filtering are computed in real time by a low-complexity polyphase filter system. In addition, the proposed technique is independent of the modulation being used, making it possible to extend the MM concept to higher order and/or non-constant amplitude constellations. Also, any number of base-system blocks can be added in a serial cascade, thereby making it possible to almost eliminate the back-off, even at very low filter roll-off values. Simulation results for M-ary constellations with M up to 64 are presented, proving the efficiency of the proposed scheme.

Like the strategy followed in LUT-MM schemes [4, 32, 34, 108], increased sensitivity to noise due to MM is compensated for by using error control coding. In forward error correction arrangements with iterative soft-decoding at the receiver successful decoding is highly dependent on the initial receiver's soft-decoding estimates. We present two efficient techniques for improving initial estimates that considerably minimizes noise sensitivity due to MPMM.

This chapter is mainly based on [36], [37], [35] and [38].

4.1 MULTIRATE SIGNAL PROCESSING OVERVIEW

Multirate systems [3, 44, 76] are systems in which signal samples are processed and manipulated at different rates, at different points of the configuration. There are many applications where the signal at a given sampling rate needs to be converted into another signal with a different sampling rate. For instance, it is very common for digital modems to use a 4-times oversampled version of the data to reduce both the distortion caused by the analog filtering and the cost of the analog filter following the DAC [44]. Accordingly, a Nyquist digital pulse shaping filter is usually designed to operate at 4 or 8 samples per symbol, i.e. with $L = 4$ or 8 in figures 1 and 10.

4.1.1 *Resamplers*

The two basic components in sampling rate alteration are the *up-sampler* and the *down-sampler*, whose most common symbols are shown in figure 19.

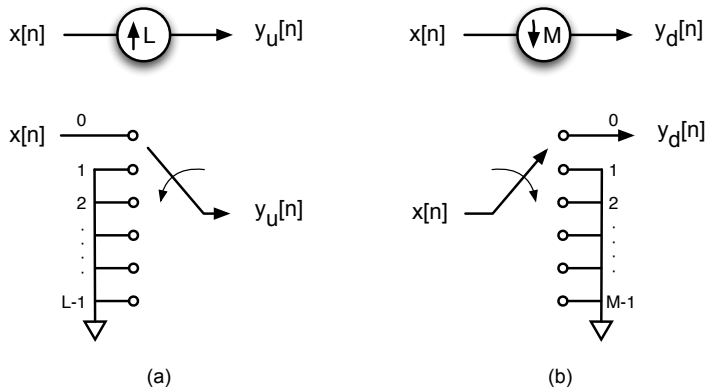


Figure 19: Symbols representing the (a) up-sampling and (b) down-sampling operations.

The two operations are duals and the systems that employ them for sample rate changes are also seen to be duals, as long as there is no *aliasing* in the down-sampling procedure. This can easily be proved by building on their time-domain or frequency-domain characterizations [3, 76].

The up-sampling operation with an up-sampling factor $L \in \mathbb{N}$ consists of inserting $L-1$ zero-valued samples between two consecutive samples of the input signal $x[n]$ according to the relation

$$y_u[n] = \begin{cases} x[n/L] & \Leftarrow \quad n = 0, \pm L, \pm 2L, \dots \\ 0 & \Leftarrow \quad \text{otherwise} \end{cases} . \tag{4.1}$$

In the transform domain we get

$$Y_u(z) = \sum_{n=-\infty}^{\infty} y[n]\, z^{-n} = \sum_{\substack{n=-\infty \\ n \bmod L = 0}}^{\infty} x\left[\frac{n}{L}\right] z^{-n} = \sum_{k=-\infty}^{\infty} x[k] \left(z^L\right)^{-k} , \tag{4.2}$$

i.e. it follows that

$$Y_u(z) = X\left(z^L\right) , \tag{4.3}$$

or equivalently, using the *discrete-time Fourier Transform* (DTFT), it can be expressed as

$$Y_u\left(e^{j\omega}\right) = X\left(e^{j\omega L}\right) \ . \tag{4.4}$$

The up-sampling effect of stretching the discrete-time axis induces an expected compression in frequency with an L-fold replication of the input spectrum in the interval $[-\pi, \pi]$. A time and spectral description example in a 1-to-3 up-sampling process is shown in figure 20.

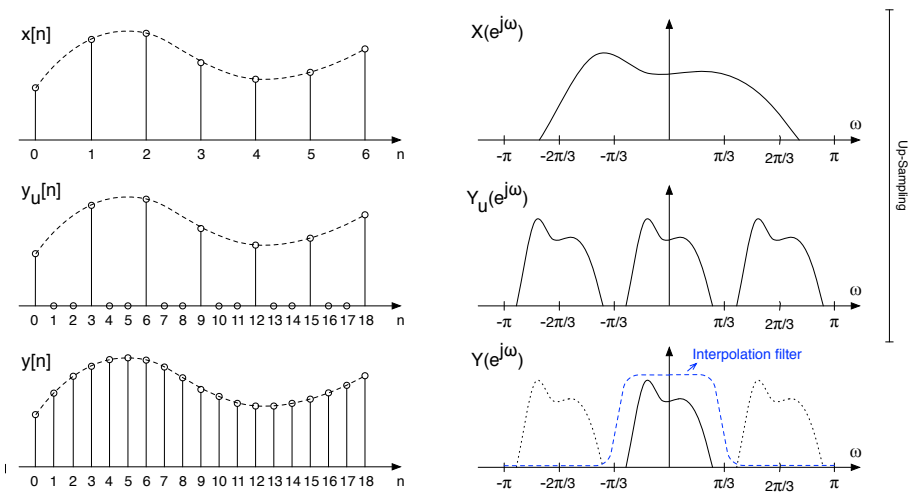


Figure 20: Time and spectral illustration of the up-sampling process for $L = 3$.

In practice, following the up-sampling, a filter limits the band of interest by removing the unnecessary spectral replicas. This is consistent in the time-domain with the zero-valued samples being replaced with interpolated values. This process, called *interpolation*, is sketched in the bottom graphs of figure 20, with the resulting interpolated signal being denoted as $y[n]$.

A down-sampler, meanwhile, with down-sampling factor $M \in \mathbb{N}$ keeps every M-th sample of the input sequence and removes the $M-1$ in-between samples. The output signal is thus given by

$$y_d[n] = x[nM] \ . \tag{4.5}$$

This process is usually visualized as a two-step procedure, with the aid of an intermediate signal $x'[n]$ operating at the same clock rate as $x[n]$, where

$$x'[n] = x[n] \cdot \sum_{r=-\infty}^{\infty} \delta(n - rM) \ , \tag{4.6}$$

and,

$$y_d[n] = x'[nM] \ . \tag{4.7}$$

Expanding the *periodic impulse train*, $\sum_{r=-\infty}^{\infty} \delta(n - rM)$, in a Fourier series [82], $x'[n]$ can be expressed as

$$x'[n] = \frac{1}{M} \sum_{k=0}^{M-1} x[n]\, e^{j k \frac{2\pi}{M} n} \ . \tag{4.8}$$

Hence, its transform is simply

$$X'(z) = \frac{1}{M} \sum_{k=0}^{M-1} X\left(z e^{-j \frac{2\pi}{M} k}\right) = \frac{1}{M} \sum_{k=0}^{M-1} X\left(z W_M^k\right) \ , \tag{4.9}$$

where $W_M = e^{-j \frac{2\pi}{M}}$. The spectrum of the sampled signal $x'[n]$ is thus the sum of M replicas of the input signal spectrum spaced $\frac{2\pi}{M}$ apart. *Aliasing* occurs unless the input signal spectrum has a bandwidth lower than $\frac{\pi}{M}$ [3, 44, 76].

The spectrum characterization of the down-sampling signal $y_d[n]$ can thus be easily derived. Applying the z-transform to (4.7), we arrive at

$$Y_d(z) = \sum_{n=-\infty}^{\infty} x'[nM]\, z^{-n} = \sum_{r=-\infty}^{\infty} x'[r]\, z^{-\frac{r}{M}} = X'\left(z^{\frac{1}{M}}\right) \ . \tag{4.10}$$

By substituting equation (4.9) in (4.10) we get

$$Y_d(z) = \frac{1}{M} \sum_{k=0}^{M-1} X\left(z^{\frac{1}{M}} W_M^k\right) \ , \tag{4.11}$$

or using the DTFT,

$$Y_d\left(e^{j\omega}\right) = \frac{1}{M} \sum_{k=0}^{M-1} X\left(e^{j\left(\frac{\omega - 2\pi k}{M}\right)}\right) \ . \tag{4.12}$$

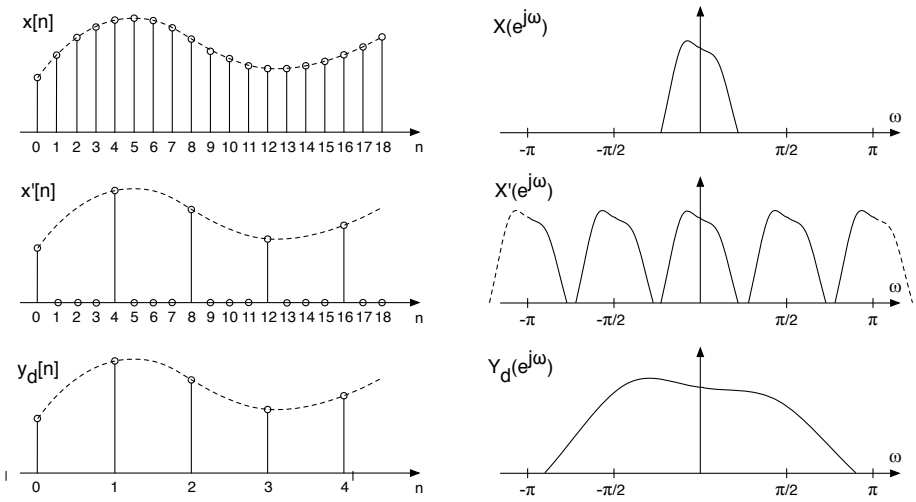


Figure 21: Time and spectral illustration of the down-sampling process for $M = 4$.

Thus the down-sampling effect of compressing the discrete-time axis induces a stretch in the frequency-domain as shown in figure 21. The down-sampling procedure can be completely reversed by performing an interpolation by M, as long there is no aliasing when down-sampling.

In order to avoid aliasing the up-sampler is usually preceded by a low-pass filter to reduce the bandwidth to $\frac{\pi}{M}$. This procedure of anti-aliasing low-pass filtering followed by down-sampling is usually called *decimation*.

4.1.2 *Noble Identities*

Building on the up- and down-sampling definitions (4.1) and (4.5), it is straightforward to conclude that equivalences of figure 22 are valid.

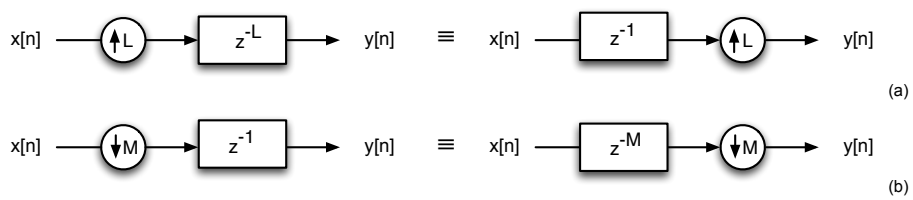


Figure 22: Equivalent delay's structures under a rate change.

In fact, focusing on figure 22.(a), it is easy to conclude that L-units of delay following a 1-to-L up-sampler is equivalent to 1-unit of delay followed by a 1-to-L up-sampler, since one clock-cycle at the input of the up-sampler spans

the same time interval as L clock cycles at its output. A similar conclusion can be drawn for case (b).

By using the previous properties, and also the fact that summing or scaling operations on sequences can commute with their sample rate change, we get the *noble identities* shown in figure 23. These define the conditions under which a filter can be pulled through the resampler.

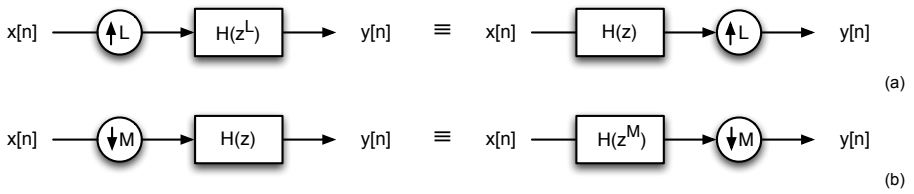


Figure 23: Noble Identities.

4.1.3 *Multirate Polyphase Signal Processing*

The noble identities are particularly useful for the efficient realization of multirate FIR filters [3, 76]. In this section we will only address the interpolator case, i.e. a FIR filter $H(z)$ following a 1-to-L up-sampler as presented in figure 24, since this is the basis of the new LUT-less approach of the concept of MM.

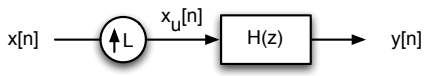


Figure 24: Multirate filter structure: interpolator.

Let us assume, without loss of generalization, a causal FIR transfer function $H(z)$ of order N:

$$H(z) = h(0) + h(1)\,z^{-1} + h(2)\,z^{-2} + \cdots + h(N)\,z^{-N}\,. \tag{4.13}$$

Given any integer number $L \leqslant N+1$, expression (4.13) can be rewritten as

$$\begin{aligned} H(z) = &\left(h[0] + h[L]\,z^{-L} + h[2L]\,z^{-2L} + \cdots\right) \\ &+ z^{-1}\left(h[1] + h[L+1]\,z^{-L} + h[2L+1]\,z^{-2L} + \cdots\right) + \cdots \\ &\cdots + z^{-(L-1)}\left(h[L-1] + h[2L-1]\,z^{-L} + h[3L-1]\,z^{-2L} + \cdots\right)\,, \end{aligned} \tag{4.14}$$

or, in a more compact and generalized form, as

$$H(z) = \sum_{i=0}^{L-1} z^{-i} E_i\left(z^L\right) , \tag{4.15}$$

where

$$E_i(z) = \sum_{n=0}^{\lfloor (N+1)/L \rfloor} h[nL + i + \lambda]\, z^{-n} , \quad 0 \leqslant i \leqslant L-1 . \tag{4.16}$$

This decomposition of $H(z)$ is usually known as *polyphase decomposition*. The parameter $\lambda \in \mathbb{Z}$ is the filter decomposition phase offset and corresponds in practice to a time-delay ($\lambda < 0$) or advance ($\lambda > 0$) of the filter impulse response.

A computationally efficient realization of the up-sampler and filter combination can thus be obtained from (4.15) and (4.16). First $H(z)$ in figure 24 is replaced by the polyphase bank as shown in figure 25.(a). Then, by using the noble identity of figure 23.(a), the upsampler can be pulled through the polyphase filters as shown in 25.(b). Note that in this realization the polyphase filters operate at a much lower rate.

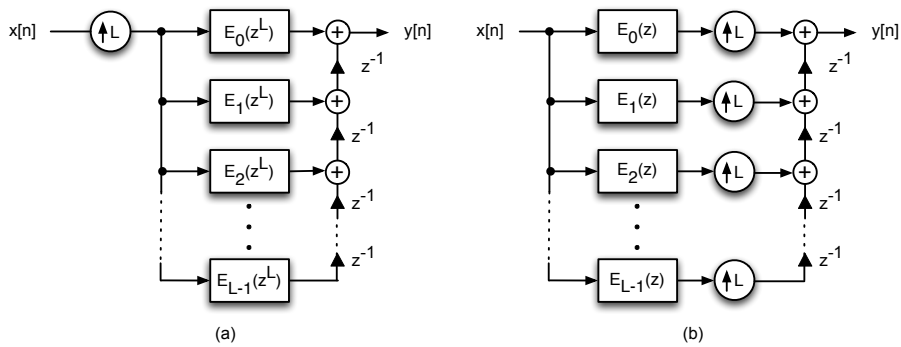


Figure 25: Interpolator's polyphase decomposition.

4.2 MULTISTAGE POLYPHASE MAGNITUDE MODULATION

It is this efficient multirate polyphase decomposition of the pulse shaping filter (joint combination of the 1-to-L up-sampler followed by filter $H(z)$) that underlies the idea of a LUT-less MM approach. The generic block diagram of a typical SC transmitter employing MM, presented in section 3.1.1, is reproduced in figure 26, but here the pulse shaping filter is displayed in its multirate polyphase efficient form.

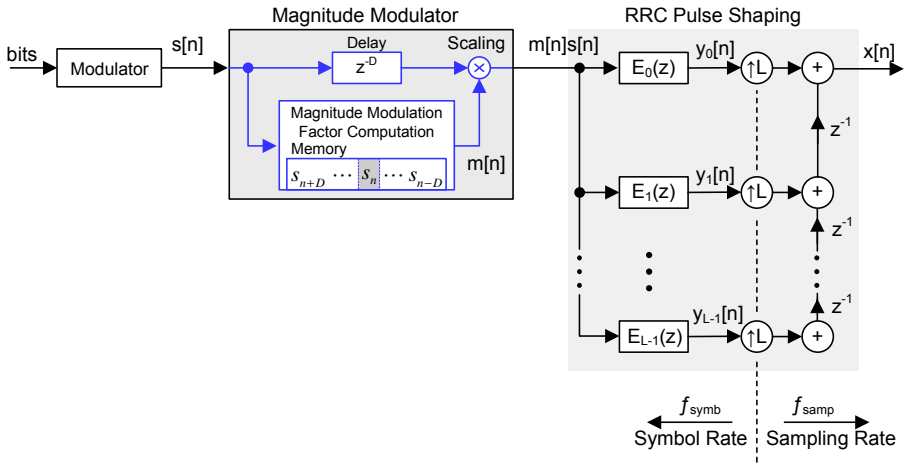


Figure 26: Generic block diagram of a typical SC transmitter system that makes use of the MM concept.

4.2.1 *The Polyphase Magnitude Modulation Principle*

In the development that follows, consider first that no MM is being used, i.e. $m[n]=1$. Let us assume, without loss of generalization, a type I linear phase FIR RRC filter [76] whose impulse response spreads over $2N+1$ symbols and which is given by:

$$H(z) = \sum_{n=0}^{2NL} h[n]\, z^{-n} \quad \text{with} \quad h[n] = h[2NL-n]\,, \tag{4.17}$$

and that the over-sampling factor L is even. Similar developments for a type II filter [76] or an odd L can easily be inferred.

According to (4.15) and (4.16), pulse shaping filtering can be performed at symbol rate using the polyphase scheme displayed in figure 26, with

$$E_i(z) = \sum_{n=0}^{2N} e_i[n]\, z^{-n} = \sum_{n=0}^{2N} h[nL+i+\lambda]\, z^{-n}\,, \tag{4.18}$$

considering the required zero padding shown in figure 28, consistent with a decomposition phase offset[1] $\lambda = -L/2$.

1 The reason for selecting $\lambda = -L/2$ is related to the center of symmetry of the RRC impulse response and the processing of the symbol to be MM. This will become clear later in this section.

The output $x[n]$ is given by one of the polyphase filters' outputs at instant $n_0 = \lfloor n/L \rfloor$, since

$$x[n] = y_{(n \bmod L)}[\lfloor n/L \rfloor] \ . \tag{4.19}$$

All outputs $y_i[n_0]$, with $i = 0, \cdots, L-1$, depend on the same set of inputs symbols $s[n']$, with $n' = n_0 - 2N, \cdots, n_0$, since

$$y_i[n] = \sum_{k=0}^{2N} e_i[k]\, s[n-k] \ . \tag{4.20}$$

The MPMM principle consists of computing the best MM coefficient $m[n_0]$ to apply to symbol $s[n_0]$, so as to guarantee that

$$|x[n]| \leqslant A \ , \tag{4.21}$$

where A is the specified polar threshold[2].

Computation of $m[n_0]$ takes into account the RRC output samples at those sampling instants to which $s[n_0]$ most contributes. Once the energy of the impulse response $h[n]$ is centered around sample NL, $s[n_0]$ is, according to (4.19) and (4.20), the most relevant symbol to the RRC output excursion $x[n]$ during the interval $\left[\left(n_0 + N - \frac{1}{2}\right)L;\left(n_0 + N + \frac{1}{2}\right)L\right[$, as shown in figure 27. Those samples can be observed at the same time instant at the output of the polyphase filter bank, $E_i(z)$ with $i = 0, \cdots, L-1$, when considering a polyphase decomposition with offset $\lambda = -L/2$.

Based on this premise and setting $\lambda = -L/2$, we have developed the system shown in figure 29 that adjusts the amplitude of symbol $s[n_0]$ in order to control the output excursion of all polyphase filters $E_i(z)$ at sampling instant $n_0 + N$ (figure 27), so as to guarantee that

$$|y_i[n_0 + N]| \leqslant A \ . \tag{4.22}$$

2 See section 3.1.1.

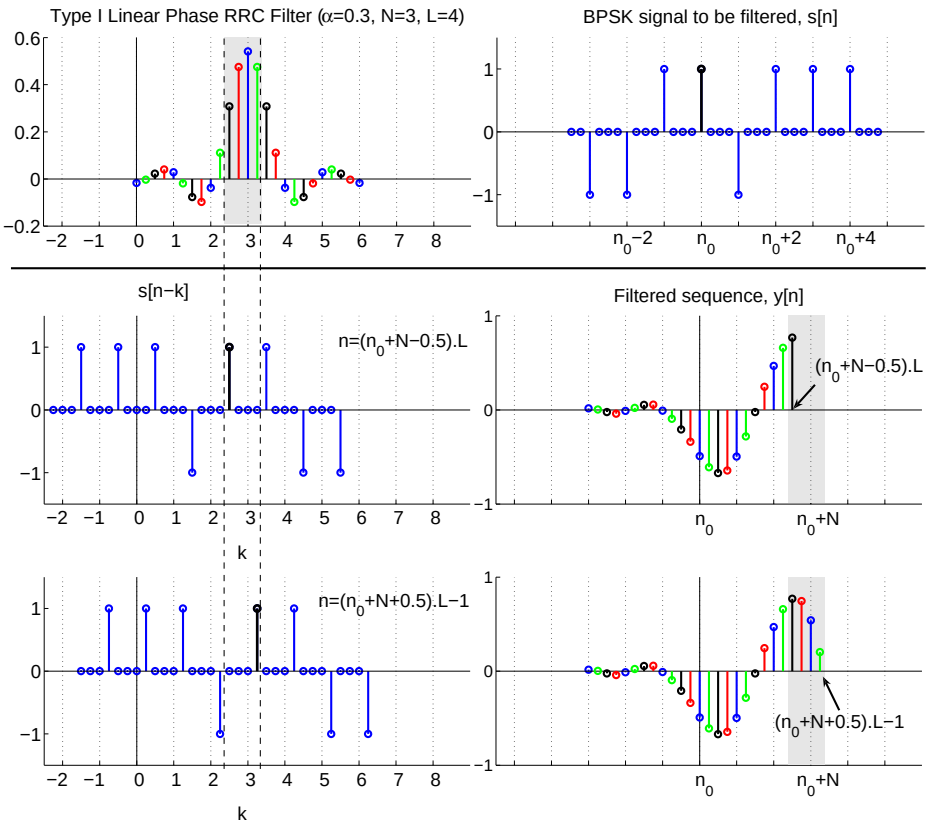


Figure 27: Controlling the output around sample n_0+N, within a range of L samples.

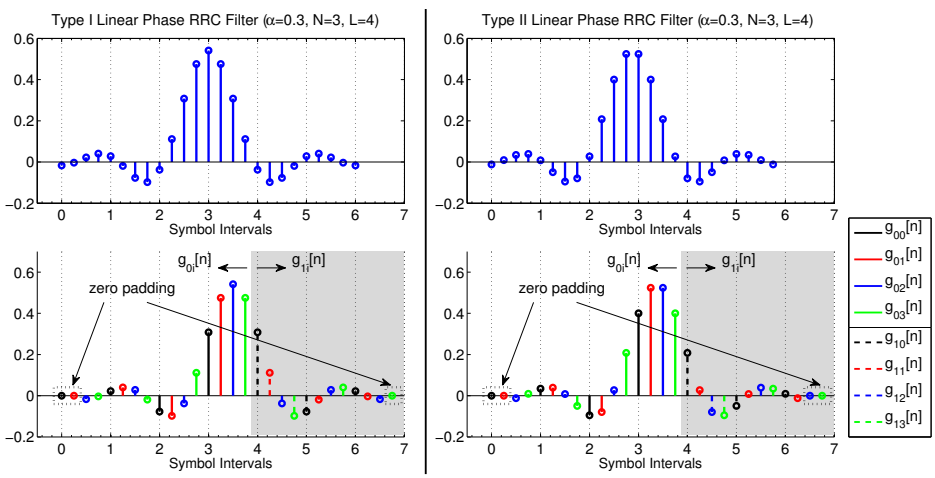


Figure 28: MPMM component filters, $g_{0i}[n]$ and $g_{1i}[n]$, for polyphase decomposition with phase-offset, $\lambda = -L/2$.

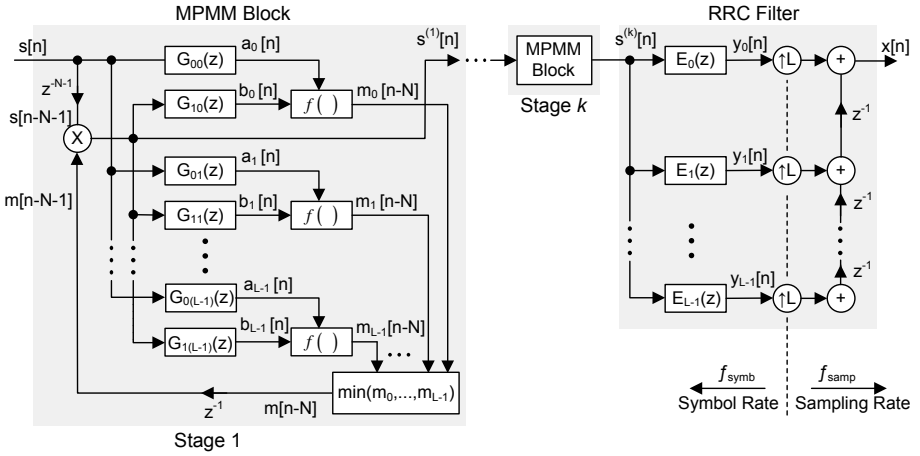


Figure 29: Multistage polyphase magnitude modulation scheme for controlling the signal excursion at the RRC output, followed by the RRC filter block.

The RRC impulse response is divided into two parts around its center of symmetry, as shown in figure 28. The polyphase filters $E_i(z)$ are consequently separated into filters $G_{0i}(z)$ and $G_{1i}(z)$, respectively on the left and right, with their impulse responses given by

$$g_{0i}[n] = \begin{cases} e_i[n] & , 0 \leqslant n \leqslant N \\ 0 & , \text{ otherwise} \end{cases} , \tag{4.23}$$

$$g_{1i}[n] = \begin{cases} e_i[n+N+1] & , 0 \leqslant n \leqslant N-1 \\ 0 & , \text{ otherwise} \end{cases} . \tag{4.24}$$

MPMM in detail...

Consider now that MM is being used. Figure 29 shows a k-stage MPMM system, followed by the polyphase representation of the RRC filter. For a better understanding and without loss of generality we will consider the case of only 1-stage and focus our attention on figure 30, which depicts a single branch of the MPMM system. When magnitude modulation is used (4.20) is written as

$$y_i[n] = \sum_{k=0}^{2N} e_i[k]\, m_i[n-k]\, s[n-k] , \tag{4.25}$$

where $m_i[n]$ is the sequence of MM coefficients which controls the peak power at the output of polyphase filter $E_i(z)$.

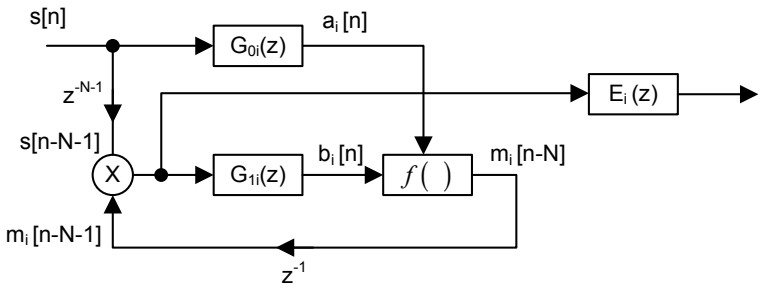


Figure 30: Detailed MPMM branch.

Basically, the system tries to anticipate the output of $E_i(z)$ at each instant n_0+N and it computes accordingly the MM factor $m_i[n_0]$ to apply to $s[n_0]$, the input sample that most contributes to the output $y_i[n_0+N]$. Each MPMM stage system therefore introduces a small delay of NT_{SYMB} during transmission.

Building on (4.25), condition (4.22) can be written as

$$\left| \sum_{k=0}^{2N} e_i[k]\, m_i[n_0+N-k]\, s[n_0+N-k] \right| \leqslant A\ . \tag{4.26}$$

When computing $m_i[n_0]$, coefficients $m_i[n_0-k]$, with $k=1,\cdots,N$, are already known (past symbols relative to $s[n_0]$). However, nothing is known about the MM values that will magnitude modulate symbols $s[n_0-q]$, with $q=-N,\cdots,-1$ (future symbols relative to $s[n_0]$). In order to avoid excessive time variation of the average power of the signal after pulse shaping, we assume that future symbols should be MM as,

$$m_i[n_0-q] \simeq m_i[n_0]\,, \quad \text{for} \quad q=-N,\cdots,-1\ . \tag{4.27}$$

Hence, considering equality in equation (4.27), condition (4.26) can rewritten as

$$\left| \begin{array}{l} m_i[n_0] \sum\limits_{k=0}^{N} e_i[k]\, s[n_0+N-k] \\ \qquad + \sum\limits_{k=N+1}^{2N} e_i[k]\, m_i[n_0+N-k]\, s[n_0+N-k] \end{array} \right| \leqslant A\ , \tag{4.28}$$

and so, using definitions (4.23) and (4.24) it follows that

$$\left| \begin{array}{l} m_i[n_0] \sum\limits_{k=0}^{N} g_{0i}[k]\, s[n_0+N-k] \\ \qquad + \sum\limits_{k=0}^{N-1} g_{1i}[k]\, m_i[n_0-k-1]\, s[n_0-k-1] \end{array} \right| \leqslant A\ . \tag{4.29}$$

Let $a_i[n]$ and $b_i[n]$ denote the complex signals at the output of filters $G_{0i}(z)$ and $G_{1i}(z)$ as shown in figure 30. Examining (4.29) and considering convolution definition, it is straightforward to conclude that

$$\sum_{k=0}^{N} g_{0i}[k]\, s[n_0 + N - k] = a_i[n_0 + N] \ , \tag{4.30}$$

and,

$$\sum_{k=0}^{N-1} g_{1i}[k]\, m_i[n_0 - k - 1]\, s[n_0 - k - 1] = b_i[n_0 + N] \ . \tag{4.31}$$

As a result, condition (4.29) can be expressed as

$$|m_i[n_0]\, a_i[n_0 + N] + b_i[n_0 + N]| \leqslant A \ . \tag{4.32}$$

A non-negative $f_{MM}(\cdot)$ function is defined so as to guarantee (4.32) with $m_i[n_0]$ being computed as

$$m_i[n_0] = f_{MM}(A, a_i[n_0 + N]\,, b_i[n_0 + N]) \ , \tag{4.33}$$

where

$$f_{MM}(A, a, b) = \begin{cases} 1 & , |a+b| \leqslant A \\ \frac{-\Re\{ab^*\} + \sqrt{\Re\{ab^*\}^2 - |a|^2(|b|^2 - A^2)}}{|a|^2} & , |a+b| > A \end{cases} \ , \tag{4.34}$$

for $A \in \mathbb{R}^+$ and $a,\ b \in \mathbb{C}$.

According to the following proposition, $f_{MM}(\cdot)$ provides the maximum value for $m_i[n_0] \in \,]0; 1]$ that satisfies (4.32), as long as

$$|b_i[n_0 + N]| \leqslant A \ . \tag{4.35}$$

This condition is always verified in practice. In fact, note that signal $b_i[n_0 + N]$ is the contribution to $y_i[n_0 + N]$ from the past symbols already attenuated by $m_i[n_0 - k]$, with $k = 1, \cdots, N$, in order to guarantee proper envelope control. Having $|b_i[n_0 + N]| > A$ would thus be nonsense, meaning that MM system was not functioning properly.

Proposition 4.1. *Given* $a, b \in \mathbb{C}$, $A \in \mathbb{R}^+$, *if* $|b| < A$ *then there exists an* $m \in \,]0; 1]$ *that satisfies*

$$|ma + b| \leqslant A \tag{4.36}$$

and its maximum value is given by 4.34.

Proof. We start by proving that if $|b| < A$ an $m \in \mathbb{R}^+$ solution exists that meets condition (4.36).

According to vector algebra properties

$$|ma + b| \leqslant |ma| + |b| \ , \tag{4.37}$$

and so, condition (4.36) will also be satisfied for any m that assure

$$|ma| + |b| \leqslant A \ , \tag{4.38}$$

whose solution can be computed as

$$|m| \leqslant \frac{A - |b|}{|a|} \ , \quad \text{with } |a| \neq 0 \ . \tag{4.39}$$

Since we are assuming $|b| < A$, the righ-hand side of (4.39) is always positive which proves the existence of an $m \in \mathbb{R}^+$ that satisfies (4.36).

In order to prove that the maximum value for $m \in]0;1]$ that meets (4.36) is given by (4.34), two different cases must be considered:

- If $|a + b| \leqslant A$, it is straightforward to conclude that maximum value for $m \in]0;1]$ that satisfies (4.36) is

$$m = 1 \ ; \tag{4.40}$$

- If $|a + b| > A$, it means that condition (4.36) is violated for $m = 1$.

  Hence, according to (4.39), solution on the maximum value of m will belong to $]0;1[$ and it will satisfy condition (4.36) with equality. It can be so computed as

$$\begin{aligned} |m\,a+b| = A &\Leftrightarrow (m\,a + b)(m\,a + b)^* = A^2 \\ &\Leftrightarrow |a|^2 m^2 + 2\mathrm{Re}\{ab^*\}\,m + |b|^2 - A^2 = 0 \ . \end{aligned} \tag{4.41}$$

  Both solutions of the quadratic equation (4.41) are real, as proved, since we are assuming $|b| < A$. Because we look for the maximum, we select the highest root which is given by

$$m = \frac{-\Re\{ab^*\} + \sqrt{\Re\{ab^*\}^2 - |a|^2(|b|^2 - A^2)}}{|a|^2} \ , \tag{4.42}$$

which is positive attending to

$$|b| < A \quad \Rightarrow \quad \sqrt{\Re\{ab^*\}^2 - |a|^2(|b|^2 - A^2)} > \Re\{ab^*\}\,. \tag{4.43}$$

□

The procedure just described computes the MM factor $m_i[n]$ that multiplies each symbol $s[n]$ in order to limit the excursion at the output of a particular filter $E_i(z)$. However, $m[n]$ must be unique and has to guarantee that $|y_i[n]| \leqslant A$ at the output of all filters $E_i(z)$ (with $i = 0, \cdots, L-1$). We have solved this problem by multiplying each symbol $s[n]$ by the most restricted factor, i.e. by

$$m[n] = \min_{i=0,\cdots,L-1} (m_i[n])\ . \tag{4.44}$$

4.2.2 *MM Factors Accuracy Improvement*

In (4.28), equality in condition (4.27) is assumed to be true in order to compute $m[n_0]$. So what happens if the next value to be computed, $m[n_0+1]$, differs from $m[n_0]$?

When $m[n_0+1] \leqslant m[n_0]$, the condition $|y_i[n_0+N]| \leqslant A$ still holds because we are applying a more restricted factor to the next symbol $s[n_0+1]$.

If $m[n_0+1] > m[n_0]$ then the factor to apply to $s[n_0+1]$ that guarantees $|y_i[n_0+N+1]| \leqslant A$ (for all i) can be relaxed, but it will no longer guarantee that $|y_i[n_0+N]| \leqslant A$. However, since symbol $s[n_0+1]$ is more relevant to the output $y_i[n_0+N+1]$ than to $y_i[n_0+N]$, there must be a compromise for $m[n_0+1]$ that does not decrease $y_i[n_0+N+1]$ too much (which is undesirable) and still guarantees that restriction $|y_i[n_0+N]| \leqslant A$ is only slightly violated. A simple procedure to accomplish this was found by using the following time variant filter:

$$m[n+1] > m[n] \ \Rightarrow\ m'[n+1] = \frac{m[n+1] + m[n]}{2} \tag{4.45}$$

Different weight averaging functions have also been tested, with unequal contributions of $m[n_0+1]$ and $m[n_0]$. As an example, tests were conducted for non-constant amplitude constellations where the relative amplitudes of symbols $s[n_0+1]$ and $s[n_0]$ were taken into account on the weights' stipulation.

However, on performing simulations for different M-ary constellations the simple arithmetic mean (4.45) proved to be the most effective of all the averages considered.

4.2.3 *Multistage PMM*

One of the major advantages of the proposed polyphase MM scheme is that the system design depends only on the RRC impulse response. There are no constraints on the input type signal which makes this new proposed technique independent of the constellation being used. The system adjusts the amplitude of the input symbols regardless of their type, which means that this procedure can also be applied to a sequence that has already been processed by MM. Each stage incrementally adjusts the symbols' amplitudes, in order to achieve a better control of the signal's excursion at the RRC's output. The back-off to apply to the input signal can therefore become almost negligible by using a cascade of a few MPMM blocks, as shown in figure 29.

4.3 MPMM SIMULATION RESULTS FOR PAPR AND BACK-OFF

This section reports reduction gains obtained with the MPMM scheme in the PAPR of the band limited signal $x[n]$ (see figure 29) and in the required back-off to drive the HPA close to saturation.

The MPMM technique was evaluated for different M-ary constellations. Unlike the LUT-MM algorithm, there is no restriction on applying the method to constellations with $M \geqslant 16$. We choose to present some of the obtained results through simulation for the 8-PSK, 16-APSK (with $\gamma_1 = 3.15$ [19]) and 32-APSK (with $\gamma_1 = 2.84$ and $\gamma_2 = 5.27$ [19]) constellations that find particular application in satellite communications, mainly due to their low $\text{PAPR}_{\text{Const}}$. The performance of the normalized 16-QAM and 64-QAM constellations under MPMM is also evaluated. Their study is of particular interest: the first, in order to compare with the 16-APSK case when MM is used[3]; the second, to assess the MPMM's performance for constellations with high number of symbols.

With the exception of the 64-QAM system, the RRC pulse shaping filters used in the simulation were designed with an oversampling $L = 16$ and a delay of $N = 7$ symbols [44, 108], where the filter's coefficients were optimized in order to ensure an ISI$< 1\%$ [11, 44]. The 64-QAM simulation considers RRC filters designed with $L = 4$ and $N = 18$.

Figures 31 and 32 show the reduction in the required back-off obtained with the MPMM method, and the PAPR of the MM signal when a different number of stages are considered.

In figure 31 the MPMM performance is compared among different constellations for the demanding case of sharp roll-off RRC filtering with $\alpha = 0.2$ used in DVB-S2 [19], and also $\alpha = 0.35$ for comparison with figures 17 and 18 (LUT-MM performance). For a better comparison with the LUT-MM technique, the best results for a memory size of $N = 5$ from the different approaches proposed in

3 Although presenting a slightly worse BER performance, the 16-APSK transmission is usually preferred to 16-QAM due to its low $\text{PAPR}_{\text{Const}}$. We want to see whether that advantage is maintained when MM is used to control the envelope of the transmitted signal.

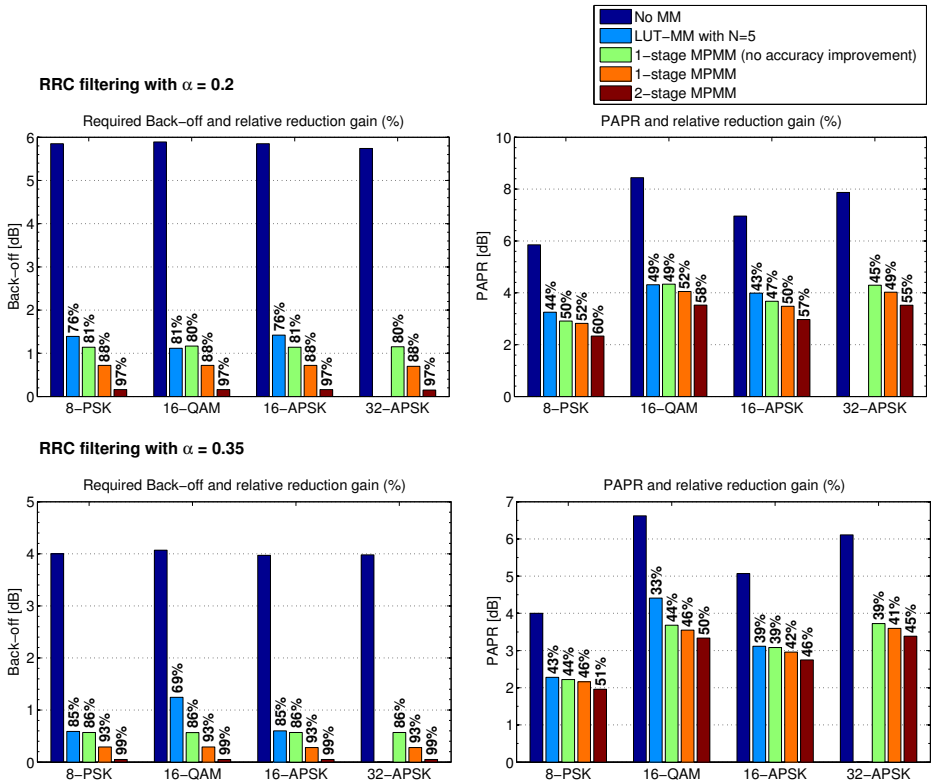


Figure 31: Back-off and PAPR gains of MPMM applied to different constellations when performing bandwidth pulse shaping with $\alpha = 0.2$ and 0.35.

section 3.2 are included in the figure. In addition, the MPMM behavior is also compared with the 1-stage configuration without improving the accuracy of the MM factors[4] through (4.45).

MPMM clearly outperforms the LUT-MM technique. Even without such accuracy improvement a reduction in the required back-off of about 80% is obtained for just 1-stage. This reduction rises to the $88\% - 93\%$ range when (4.45) is used, which shows the effectiveness of the simple procedure proposed. For 2-stages the need to use back-off almost vanishes, with reduction gains always above 95% (similar results obtained for the different tested constellations for an even more demanding example of RRC filtering with $\alpha = 0.1$).

Figure 31 also shows that MPMM performs identically for the different constellations. This behavior was expected, since the maximum excursions above the desired limit, A, are mainly due to outer-circle symbol combinations. Consequently the system behavior should be the same in such cases, as was

4 Whenever reference is made to the MPMM technique, unless otherwise said, it is assumed that accuracy of the MM factors through (4.45) is used.

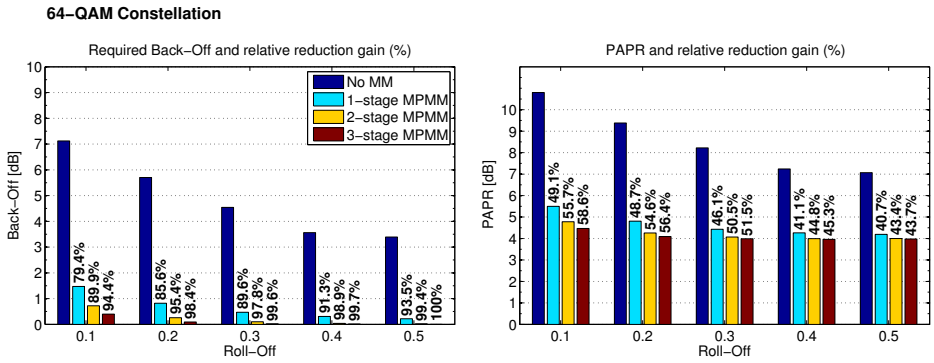


Figure 32: Back-off and PAPR gains of k-stage MPMM, with $k = 1, \cdots, 3$, applied to 64-QAM RRC filtered signal for different roll-offs.

found. This important fact demonstrates that this technique is independent of the modulation being used, as previously claimed.

The previous conclusion is emphasized by analysis of the MPMM's performance results for a 64-QAM constellation (figure 32). We observe that MPMM performs equally well for this high number of symbols. The figure plots MPMM's performance for bandwidth pulse shaping with different values of roll-off in the range $[0.1; 0.5]$. Results confirm that the MPMM method performs very well even for very stringent roll-offs, and that with just 3-stages the need to use back-off can be completely eliminated for $\alpha \geqslant 0.2$.

With respect to the PAPR, both figures show that its reduction is lower than back-off reduction. Hence, the expected fall in average transmitted power due to the MM procedure is much lower than the reduction in output peak power, as was recently shown analytically[5] [39]. This behavior is desirable. In fact, the small decrease in average transmitted power due to MM justifies the small penalty on BER, as will be seen in the next section.

Note also that the reduction in PAPR is less for high roll-off values, although the need to use back-off is completely eliminated. This is easily understandable because the need to adjust the symbol's amplitudes becomes weaker for less stringent roll-offs, and therefore the average power does not fall too much.

When comparing, under MPMM envelope control, the PAPR of a 16-QAM transmitted signal with a 16-APSK one, we observe that they exhibit a PAPR of the same order, and so the aforementioned advantage of 16-APSK having a much lower $\text{PAPR}_{\text{Const}}$ disappears. This is confirmed by the fact that a 16-QAM transmission using MM still exhibits a better BER performance than a 16-APSK constellation, as is shown in section 4.4. This higher PAPR gain in dBs for the 16-QAM occurs because only its 4 outer constellation symbols suffer more severe

5 This analytical analysis is addressed in the next chapter.

distortion due to MM, which leads to a smaller fall in the average transmitted power.

4.4 IMPROVED MPMM DETECTION AND BER PERFORMANCE

The BER performance of the proposed MPMM technique was evaluated for the AWGN channel assuming the linear transmission (ideal HPA operation) of an $\alpha = 0.2$ RRC pulse shaped signal. Forward error correction with soft-decision decoding was used, considering the $(1248, 832)$ and $(2016, 1680)$ LDPC codes defined in the WiMAX 802.16e standard [49]. Decoding was performed using the logarithmic SPA algorithm described in section 2.4.2, following a *horizontal schedule* [29, 100] approach and considering a maximum number of 100 iterations per decoded word.

4.4.1 *BER performance of an MM's blind receiver*

In section 3.2.5, when analyzing the BER performance of the LUT-MM technique we were assuming that the receiver was MM's blind, i.e. the receiver operates normally as if no MM was previously applied at the transmitter's side. Due to the distortion caused by the MM procedure to the transmitted symbols, some increased sensitivity to noise was observed. Hopefully, this loss was well overcompensated for by the reduction gain obtained in the required back-off.

MPMM exerts a fine control of the signal's envelope excursion and so a worsening of sensitivity to noise should be expected. Figure 33 plots the BER performance of an MM's blind receiver when MPMM is used in the transmitter. It shows a serious BER degradation for M-ary constellations with $M \geqslant 32$, even for a 1-stage MPMM configuration. The observed loss[6], $L_{@BER}$, kills the G_{BO}, making MPMM a technique of limited interest.

For smaller constellation size the net back-off gain is clearly positive. However, interestingly, we observe a higher BER degradation for 16-APSK transmission than for 16-QAM. As both constellations are the same size ($M = 16$), it is straightforward to conclude that sensitivity to noise due to MM depends on constellation geometry. In fact, non-constant amplitude constellations with low PAPR or with aligned symbols (same phase) are more prone to noise. This is easily understandable from the 16-APSK and 16-QAM MPMM symbols' displacement and their corresponding *pdf*, shown in figure 34. Here, we can observe for the 16-APSK that all outer symbols suffer severe distortion and that a few of them (those with phase $\phi = \pi/4$) are dangerously close to the inner ones, while for 16-QAM only the smaller set of 4 outer symbols are more severely distorted, but they are still much further away from their neighbors. In addition, note also that the outer 16-QAM symbols are less distorted than

6 Refer to the net back-off gain (net_G_{BO}) definition given in (3.34).

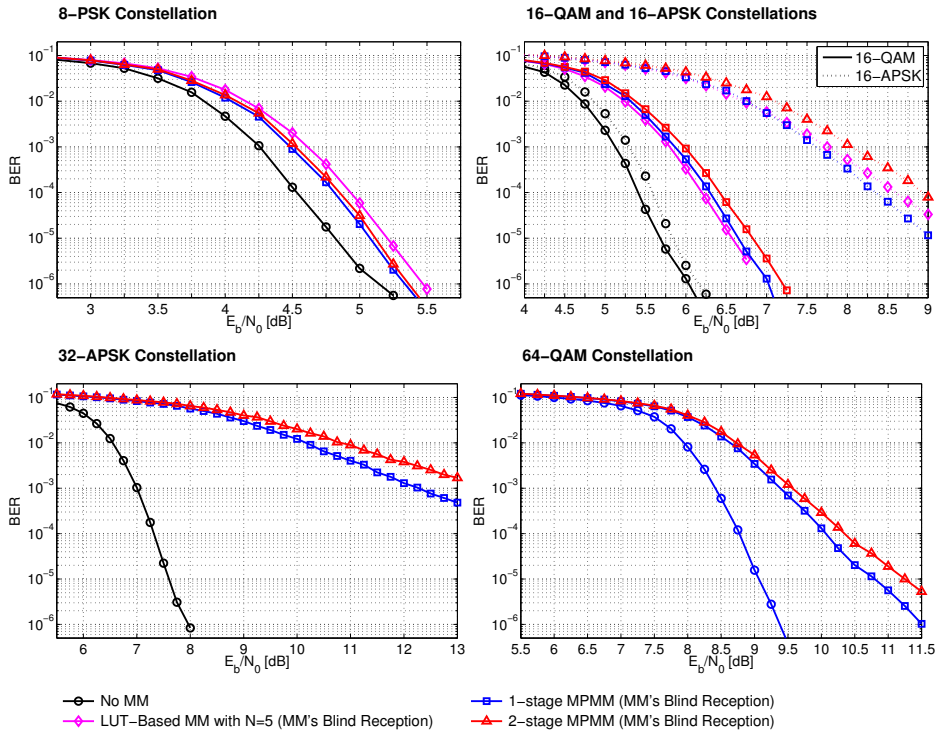


Figure 33: Soft-decoding BER performance for MM's blind reception of MPMM signals with respect to different constellations, when performing bandwidth pulse shaping with $\alpha = 0.2$ and using a FEC system based on the $(1248, 832)$ LDPC code [49].

the 16-APSK outer ones, as can be concluded from the analysis of their *pdf* distributions[7].

The same reasoning explains the severe BER degradation observed for higher order constellations. When computing initial soft-decoding estimates, an MM's blind receiver will easily commit an error (wrong estimation), even under the ideal conditions of an error free transmission environment. As the constellation symbols are close together and because of the distortion introduced by the MM procedure to control the signal's envelope excursion, an MM symbol, say $s_{MM}[n] = m[n]\, s[n]$, will mislead the receiver whenever $s_{MM}[n]$ is closer to any constellation symbol that differs from $s[n]$.

7 The *pdf* distributions of the MPMM coefficients are analyzed in the next chapter.

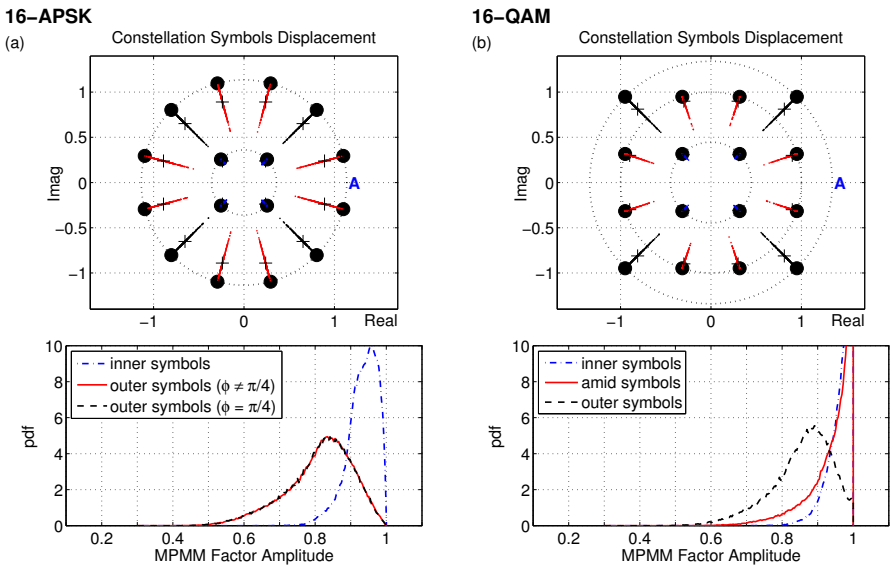


Figure 34: MPMM method applied to (a) 16-APSK and (b) 16-QAM transmission with $\alpha = 0.2$ RRC filtering. Scatter plot and pdf distributions of the symbols' displacement in a 2-stage system (• original symbol's positions; + MPMM symbols' average positions).

4.4.2 *Detection with an Estimated MM Average Constellation*

The correct estimation of a priori code bit LLRs, based on the received information from channel, plays a crucial role in the performance of a soft-decision iterative decoding algorithm like the SPA.

Let $\mathcal{X}$ denote the set of constellation symbols of a M-ary complex constellation, with $M = 2^l$ and $l \in \mathbb{N}$. The M-ary symbol at the modulator's output is represented by $s = \mathcal{F}(c_0, c_1, \cdots, c_{l-1})$, where $\mathcal{F}$ is the function that maps code bits[8] c_i, with $i = 0, \cdots, l-1$, to constellation symbols. The constellation is assumed to be normalized, i.e. with unitary average power. Considering transmission over the AWGN channel, the complex received sample at the output of the receiver matched filter may be expressed as

$$r = s + n_z \ , \tag{4.46}$$

where $n_z \in \mathbb{C}$ is a zero mean random variable representing the complex AWGN noise, with variance

$$\sigma^2_{n_z} = E\{n_z n_z^*\} \ . \tag{4.47}$$

8 Sequence of bits at channel encoder's output.

As defined in (2.95), the LLR of a code bit c_i given the received word r is

$$\mathrm{LLR}(c_i\,|r\,) = \ln \frac{\Pr(c_i = 0\,|r\,)}{\Pr(c_i = 1\,|r\,)}\ . \tag{4.48}$$

Assuming perfect knowledge of the channel, this can be computed as [110]

$$\mathrm{LLR}(c_i\,|r\,) = \ln \left[\frac{\displaystyle\sum_{\{s\in\mathcal{X}:c_i=0\}} \exp\left(-\frac{|r-s|^2}{\sigma_{n_z}^2}\right)}{\displaystyle\sum_{\{s\in\mathcal{X}:c_i=1\}} \exp\left(-\frac{|r-s|^2}{\sigma_{n_z}^2}\right)} \right] \tag{4.49}$$

where $\{s \in \mathcal{X} : c_i = b\}$, with $b \in \{0, 1\}$, is the sub-set of constellation symbols having code bit $c_i = b$.

The critical loss of performance observed in figure 33, especially for constellations with $M \geqslant 32$, is caused by an MM's blind receiver computing (4.49) based on $\mathcal{X}$, i.e. the original constellation symbols' position.

However, the initial receiver's soft-decoding estimates should take into account the symbol displacement statistics due to MM. A simple way to do this without changing the existing receivers is to compute those estimates based on the expected average MM constellation (EAC) [36, 39].

Let $\overline{s}_{i_{MM}}$ denotes the average position of constellation symbol $s_i \in \mathcal{X}$ after being MM, i.e.

$$\overline{s}_{i_{MM}} = s_i E\{m\,|s_i\,\} = s_i \int_0^1 m\, p(m\,|s_i\,)\, dm\ , \tag{4.50}$$

where $p(m\,|s_i\,)$ is the *pdf* of the MM factors given that symbol s_i was transmitted. The EAC, $\overline{\mathcal{X}}_{MM}$, is simply the set of all symbols $\overline{s}_{i_{MM}}$, with $i = 0, \cdots, M-1$, computed according to (4.50). This set can be easily obtained a priori through simulation and so it can be known by the receiver. As an example, in figure 34 we plot the MPMM EAC constellation for 16-APSK and 16-QAM transmission with $\alpha = 0.2$ RRC filtering.

The initial receiver's soft decoding estimates will therefore be computed as

$$\mathrm{LLR}(c_i\,|r\,) = \ln \left[\frac{\displaystyle\sum_{\{\overline{s}_{MM}\in\overline{\mathcal{X}}_{MM}:c_i=0\}} \exp\left(-\frac{|r-\overline{s}_{MM}|^2}{\sigma_{n_z}^2}\right)}{\displaystyle\sum_{\{\overline{s}_{MM}\in\overline{\mathcal{X}}_{MM}:c_i=1\}} \exp\left(-\frac{|r-\overline{s}_{MM}|^2}{\sigma_{n_z}^2}\right)} \right]\ . \tag{4.51}$$

Figure 35 displays the BER performance of the MPMM scheme when using the EAC for the initial soft decoding estimates. The other conditions are the

same as those in figure 33, i.e. transmission over the AWGN channel and a FEC system based on the $(1248, 832)$ LDPC WiMAX802.15e code [49].

A marked improvement in BER performance can be observed. Although the selected code rate is only 2/3, we observe that it efficiently offsets the increased noise sensitivity due to MM. Even in the worst scenario, the 32-APSK with a 2-stage MPMM configuration, the BER performance loss with respect to non-MM case at BER$= 10^{-6}$ is only $L_{@10^{-6}} = 1.6$dB. Since the back-off reduction is $G_{BO} = 5.6$dB, the power efficiency of the system is improved by more than 4dB. For 8-PSK the overall back-off gain (net_G_{BO}) reaches 5.4dB. Hence, this simple technique, which does not require any changes to existing receivers, is extremely efficient.

Figure 36 illustrates what happens when a much lower rate code is used. The BER performance for the $(2016, 1680)$ LDPC WiMAX802.15e code [49] is compared with results just presented for the $(1248, 832)$ code. By providing good estimates based on the EAC for the soft-decoder, although using different codes, the performance of the MPMM technique is similar for 8-PSK and only

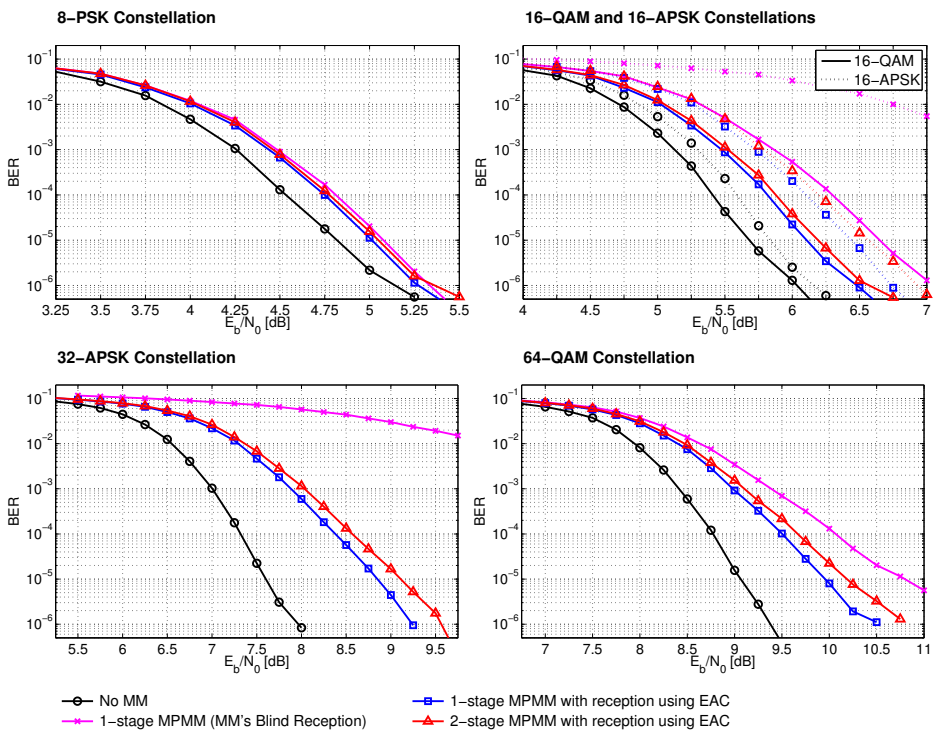


Figure 35: Soft-decoding BER performance of MPMM applied to different constellations when using the MM EAC for the initial soft decoding estimates. Tested system uses RRC pulse shaping with $\alpha = 0.2$, and FEC based on the $(1248, 832)$ LDPC code [49].

a slight degradation in BER performance is observed for 16-QAM when the weaker code is used. Hence, for constellations with $M \leqslant 16$, the EAC and the system's FEC can efficiently compensate for the increased sensitivity to noise resulting from MM. This means that no additional error control protection is required, so the information rate and the quality of service links are preserved.

However, for M-ary constellations with $M \geqslant 32$ we observe a marked BER performance degradation with net_G_{BO} vanishing. As will be shown in section 5.3.2, for higher order constellations, MM may cause a serious SER because of the constellation symbols' close proximity. Here the 5/6 rate code lacks the capacity to compensate for the increased sensitivity to noise resulting from MM and so additional error protection must be provided or an improved soft-decoding scheme used.

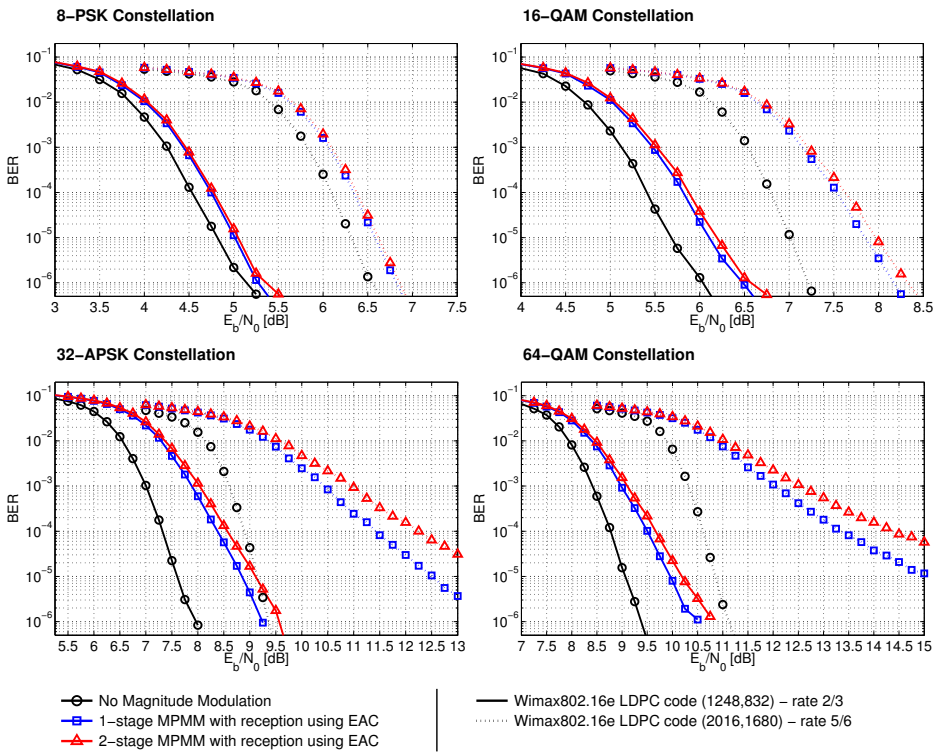


Figure 36: Comparison of soft-decoding BER performance between the $(1248, 832)$ and $(2016, 1680)$ LDPC codes [49], when using the MM EAC for the initial soft decoding estimates.

4.4.3 *Detection with Estimation of MPMM Coefficients*

Detection of an MPMM signal can be improved by proper estimation on the receiver side of the MM factors used in transmission. In this way, more accurate code bits LLRs can be provided to the receiver's soft-decoder and so a better BER performance should be expected. But correct estimation of the MM coefficients depends on the receiver's knowledge of the sequence of symbols that was sent, which the receiver does not know and is in fact looking for. Although this seams a paradox, the fact is that both the soft-decoder and the MPMM estimator can benefit from any information that each can pass to the other. Hence, the exchange of information between these two systems, analogous to the concept of iterative message-passing described in section 2.4.2, can improve decoding performance.

Figure 37 gives the block diagram of a joint design of a receiver's MPMM estimator and soft-decoder (MPMM+SD).

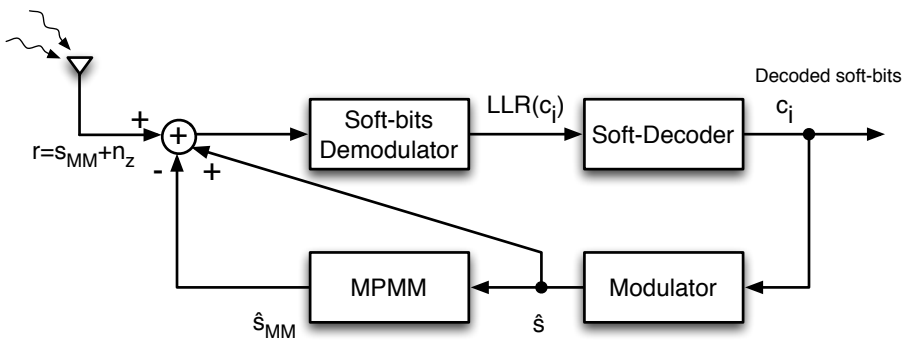


Figure 37: Joint MPMM+SD detection scheme.

Let us assume, without loss of generalization, a FEC system based on an n-length block code $\mathcal{C}$ and that each codeword

$$\mathbf{c} = [c_0,\ c_1,\ \cdots,\ c_{n-1}] \in \mathcal{C}\,, \tag{4.52}$$

is mapped into a set of constellation symbols as previously defined in section 4.4.2, i.e.

$$\mathbf{s} = \mathcal{F}(\mathbf{c}) = \left[s_0,\ s_1,\ \cdots,\ s_{(n/l)-1}\right]\,, \tag{4.53}$$

when considering an M-ary constellation with $M = 2^l$ and assuming n to be a multiple of l.

The transmitted MM codeword given by applying MPMM to $\mathbf{s}$ is represented by

$$\mathbf{s}_{MM} = \left[s_{MM_0},\ s_{MM_1},\ \cdots,\ s_{MM_{(n/l)-1}} \right] , \tag{4.54}$$

and let

$$\mathbf{r} = \left[r_0,\ r_1,\ \cdots,\ r_{(n/l)-1} \right] , \tag{4.55}$$

denote the (n/l) sampled outputs of the receiver's matched filter. According to (4.46), each complex received sample is given by

$$r_i = s_{MM_i} + n_{z_i} , \tag{4.56}$$

where n_{z_i} is a complex random variable representing the AWGN noise added to the MM transmitted symbol s_{MM_i}.

According to figure 37, based on received word $\mathbf{r}$, i.e. by setting $\mathbf{r'} = \mathbf{r}$, the *soft-bits demodulator* starts by providing code bit LLRs estimates, $\mathrm{LLR}(c_i\,|\mathbf{r'})$, to the soft-decoder.

If decoding is successful the receiver proceeds to detection of the next received word.

If decoding is unsuccessful the soft-decoder output word

$$\hat{\mathbf{c}} = [\hat{c}_0,\ \hat{c}_1,\ \cdots,\ \hat{c}_{n-1}] , \tag{4.57}$$

is presented as input to the cascade of modulator and MPMM block so as to compute new estimates of the transmitted sequence of constellation symbols,

$$\hat{\mathbf{s}} = \left[\hat{s}_0,\ \hat{s}_1,\ \cdots,\ \hat{s}_{(n/l)-1} \right] , \tag{4.58}$$

and its MPMM sequence counterpart

$$\hat{\mathbf{s}}_{MM} = \left[\hat{s}_{MM_0},\ \hat{s}_{MM_1},\ \cdots,\ \hat{s}_{MM_{(n/l)-1}} \right] . \tag{4.59}$$

In unsuccessful decoding only a few code bits of $\hat{\mathbf{c}}$ are expected to be wrong. Hence, because modulation is a memoryless procedure, estimate $\hat{\mathbf{s}}$ contains mostly valid information. The same happens with $\hat{\mathbf{s}}_{MM}$, despite MPMM not being memoryless. However, given that the influence on the MM factor that applies to a symbol $s[n_0]$ is limited to its closer neighbors, and also to the assumptions under (4.27) and (4.45), estimation $\hat{\mathbf{s}}_{MM}$ is largely reliable.

Estimates $\hat{\mathbf{s}}$ and $\hat{\mathbf{s}}_{MM}$ can thus be used to improve reliability of the information provided to the soft-bits demodulator according to

$$\mathbf{r'} = \mathbf{r} - \hat{\mathbf{s}}_{MM} + \hat{\mathbf{s}} \ , \tag{4.60}$$

i.e. we try to cancel the MM distortion added to transmitted symbol. In fact, assuming ideal estimation, i.e. $\hat{s}_{MM_i} = s_{MM_i}$ and $\hat{s}_i = s_i$, each received sample passed to the soft-bits demodulator

$$r'_i = s_{MM_i} - \hat{s}_{MM_i} + \hat{s}_i + n_{z_i} \ , \tag{4.61}$$

would reduce to

$$r'_i = s_i + n_{z_i} \ , \tag{4.62}$$

and so there would be no BER performance loss with respect to a non-MM transmission.

Decoding stops whenever a valid codeword is found, or a maximum number of feedback iterations is reached. After decoding each received word (whether successful or not), say $\mathbf{r'}_{\Delta t_n}$ over interval[9] Δt_n, and before the receiver proceeds to detect the next received word $\mathbf{r'}_{\Delta t_{n+1}}$, MPMM's internal state is updated to $S^{MPMM}_{\Delta t_n}$ by running $\hat{\mathbf{c}}_{\Delta t_n}$ through the modulator+MPMM once. At the beginning of each new iteration to decode $\mathbf{r'}_{\Delta t_{n+1}}$, MPMM's internal state is reset to $S^{MPMM}_{\Delta t_n}$.

The constellation to be used by the receiver to compute initial soft-decoding LLRs and estimates $\hat{\mathbf{s}}$ and $\hat{\mathbf{s}}_{MM}$ can be the original constellation used at the transmitter. However, since detection starts by performing soft-decoding of the received word $\mathbf{r}$ and fulfillment of this task depends critically on initial LLRs estimates (as seen in section 4.4.2), then much better results can be observed when MM EAC is used instead.

Algorithm 4.1 summarizes the joint MPMM+SD decoding procedure just described.

The MPMM+SD detection was tested for a demanding case requiring transmission of an $\alpha = 0.2$ RRC pulse shaped signal with nearly zero back-off. Soft-decoding was performed, as for the results in figure 35, using the logarithmic SPA algorithm following a horizontal schedule [29, 100] approach, but now considering only a maximum number of 50 iterations per decoded word. Initial soft decoding estimates were computed using the MM EAC and, in order to simplify decoding, a sole 1-stage MPMM system has been used at the receiver.

Figure 38 presents the BER performance of the joint MPMM+SD technique when using the $(1248, 832)$ LDPC code [49] with the maximum allowed feedback

9 Each time interval Δt_i has duration $\frac{n}{l} T_{SYMB}$, where n is the code length, l is the number of bits per constellation symbol and T_{SYMB} is the symbol's time duration.

Algorithm 4.1 Join MPMM+Soft Decoder Detection Algorithm

For a generic vector $\boldsymbol{x}$ *consider that the notation* $\boldsymbol{x}^{(i)}_{\Delta t_n}$ *denotes the value* $\boldsymbol{x}$ *during time interval* Δt_n *at iteration* (i). *During interval* Δt_n *the maximum number of admissible feedback iterations is MAX_ITER.*

STEP 0: **Initialization**

– Setup soft-decoder's input data: $\mathbf{r'}^{(0)}_{\Delta t_n} = \mathbf{r}_{\Delta t_n}$.

– Setup MPMM initial state $S^{MPMM}_{\Delta t_{n-1}}$ by running previous soft-decoder output, $\hat{\mathbf{c}}_{\Delta t_{n-1}}$, through receiver's modulator and MPMM blocks.

do {

STEP 1: **Reset MPMM's state to** $S^{MPMM}_{\Delta t_{n-1}}$

STEP 2: **Compute initial soft-decoding estimates LLR**$\left(c_i \left| \mathbf{r'}^{(i)}_{\Delta t_n} \right.\right)$ **.**

STEP 3: **Based on LLR**$\left(c_i \left| \mathbf{r'}^{(i)}_{\Delta t_n} \right.\right)$**, perform soft-decoding of** $\mathbf{r}_{\Delta t_n}$ **, generating** $\hat{\mathbf{c}}^{(i)}_{\Delta t_n}$ **.**

STEP 4: **Modulate decoded word** $\hat{\mathbf{c}}^{(i)}_{\Delta t_n}$ **in order to produce an estimate,** $\hat{\mathbf{s}}^{(i)}_{\Delta t_n}$**, of the transmitted sequence of symbols.**

STEP 5: **Magnitude modulate** $\hat{\mathbf{s}}^{(i)}_{\Delta t_n}$ **in order to produce an estimate,** $\hat{\mathbf{s}}^{(i)}_{MM_{\Delta t_n}}$**, of the transmitted sequence of MPMM symbols.**

STEP 6: **Update soft-bits demodulator input for next iteration:**

$$\mathbf{r'}^{(i+1)}_{\Delta t_n} = \mathbf{r}_{\Delta t_n} - \hat{\mathbf{s}}^{(i)}_{MM_{\Delta t_n}} + \hat{\mathbf{s}}^{(i)}_{\Delta t_n} .$$

} **while** ($\hat{\mathbf{c}}^{(i)}_{\Delta t_n}$ is an unvalid codeword) & (#iterations < MAX_ITERATIONS)

STEP 7: **Output decoded word** $\hat{\mathbf{c}}^{(i)}_{\Delta t_n}$ **from last iteration runned.**

iterations set to 3. The results refer to the 1- and 2-stage MPMM configurations used at the transmitter (Tx) for comparison with figure 35, when detection based on the EAC was studied.

Analysis of figure 38 shows that the already good results obtained earlier with the use of the MM EAC can be improved, with special relevance for constellations with $M \geqslant 32$, e.g. for 64-QAM with a Tx configuration of 1-stage MPMM, system performs only 0.5dB away from the non-MM case corresponding to a net_$G_{BO} \simeq 4.4$dB. In addition, for the case of strict fine control of the envelope excursion, i.e. for the 2-stage configuration at Tx, the BER degradation observed at the receiver is still lower than that experienced with detection based only on the EAC for the 1-stage MPMM Tx configuration. This corresponds to an additional net back-off gain of at least 0.5dB.

As seen before for large constellations ($M \geqslant 32$), the use of a detection scheme based only on the MM EAC is not sufficient to compensate for the increased sensitivity to noise resulting from MM if the channel code is not "strong" enough. Figure 39 plots the BER performance of the joint MPMM+SD technique as a

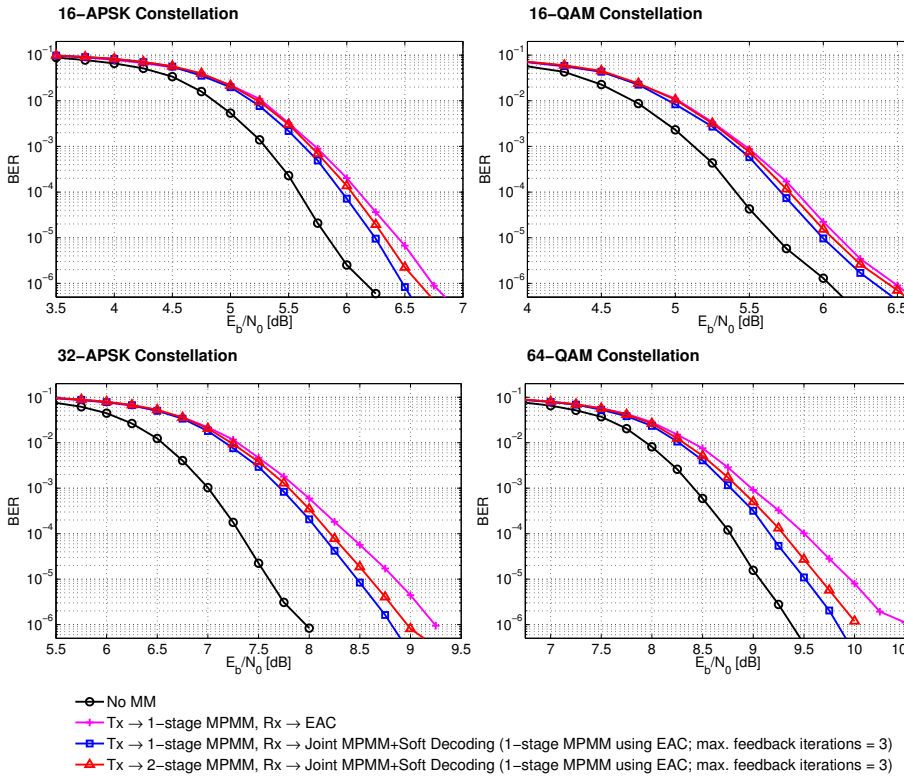


Figure 38: BER performance of joint MPMM+SD detection scheme using the $(1248, 832)$ LDPC code [49] and RRC pulse shaping with $\alpha = 0.2$.

function of the number of maximum allowed feedback iterations, when using the less powerful $(2016, 1680)$ LDPC code [49]. A 2-stage MPMM configuration is employed at Tx.

The improvement in performance for both cases is quite clear. With a maximum of only 3 feedback iterations at BER$= 10^{-5}$, this method performs only 2dB and 2.5dB away from the non-MM case for 64-QAM and 32-APSK, which corresponds to a net_G_{BO} of about 3.5dB and 3.1dB, respectively.

The results in figure 39 also show that, as expected, the decoder performs better as the number of maximum iterations allowed for the estimation of the MM factors increases, although the maximum achievable performance tends to a limit, e.g. the gain from using 7 feedback iterations over 5 is already small. In addition, the computation time also increases, and so there is a tradeoff between the number of feedback iterations, the time taken on soft-decision decoding of a received word, and also the desired performance. Experience has shown that a good tradeoff was to consider only 3 feedback iterations, while reducing the maximum number SPA soft-decoding iterations to half (i.e. 50) for our testbed.

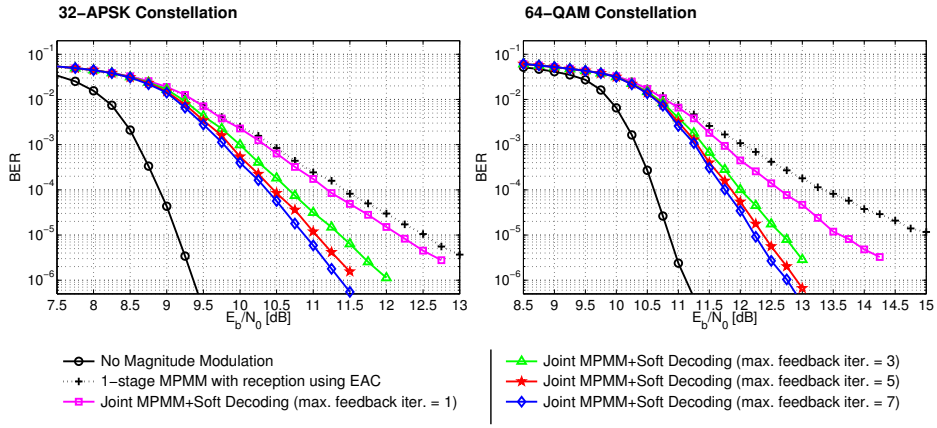


Figure 39: BER performance of joint MPMM+SD detection scheme using the $(1248, 832)$ LDPC code [49] as function of the number of allowed estimates' recomputation (max. feedback iterations) of the MM factors. Results refer to bandwidth pulse shaping with $\alpha = 0.2$ and a 2-stage MPMM Tx's configuration.

As a main conclusion, the joint MPMM+SD detection scheme makes it possible to successfully use the MPMM technique for high-order constellations even when the system employs a less powerful FEC system, with effective net_G_{BO} being experienced without the need for additional error protection.

4.5 CONCLUSION

A novel *Multistage Polyphase Magnitude Modulation* (MPMM) scheme for efficiently controlling the envelope's peak power of single carrier band limited signals has been presented in this chapter.

To compensate for the increased sensitivity to noise due to MM, two different techniques for improving the detection of the MPMM signal have also been proposed, viz.: detection based on the MM estimated average constellation (EAC); and detection with estimation of MPMM factors in the receiver. The first detection technique does not imply changes to traditional receivers and provides a notable improvement in performance with respect to MM's blind reception. The second technique combines MPMM factors estimation and soft decoding in a feedback loop scheme, and it was found to be particularly efficient for higher order constellations with $M \geqslant 32$ (more sensitive to MM distortion), especially when less powerful FEC coding is used.

The performance of the MPMM scheme was evaluated through simulation in terms of PAPR, back-off and BER. One useful measure to assess the overall performance is the net back-off gain net_G_{BO}, computed as in (3.34). Figure 40 summarizes net_G_{BO} at BER$= 10^{-5}$ for the different detection schemes, for a system employing a 2-stage MPMM configuration at transmitter, RRC pulse

shaping with $\alpha = 0.2$ and using a FEC system based on the $(1248, 832)$ LDPC code [49].

The results emphasize the high efficiency of this new proposed MPMM technique.

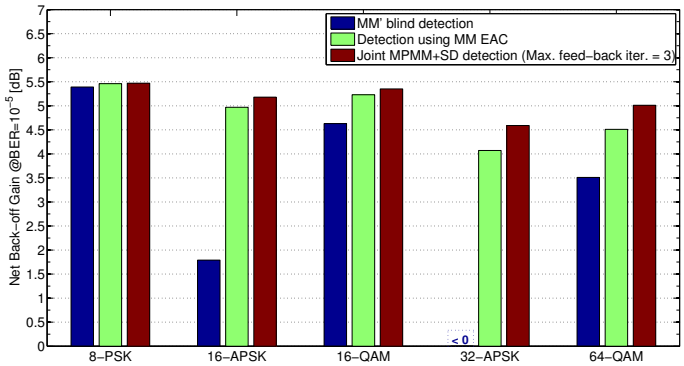


Figure 40: Net back-off gain at BER= 10^{-5} for different detection schemes, for or a system employing a 2-stage MPMM configuration at transmitter, RRC pulse shaping with $\alpha = 0.2$ and using a FEC system based on the $(1248, 832)$ LDPC code.

5

MODELING MAGNITUDE MODULATION PERFORMANCE

In this chapter we perform an analytical analysis of the transmitted signals using the peak-power reduction MM methods described in chapters 3 and 4. The statistical distributions of the MM coefficients are studied and parameterized probability density functions are provided, with typical parameters being obtained for different M-ary constellations and RRC roll-offs through simulation. Analytical average power and PAPR expressions for the transmitted MM signal are then deduced. The average drop in power due to MM is compared with the power losses observed when using other common peak-power reduction techniques, i.e. *scaling* (pure back-off applied to the signal) and *clipping*.

A typical figure of merit to evaluate the accuracy of a modulation scheme when subjected to magnitude and phase distortions is the *error vector magnitude* (EVM). The analytical study of the EVM and *symbol error rate* (SER) due to MM distortion is also addressed. Assuming a noiseless transmission, we derive expressions for the analytical computation of the EVM and bounds on the SER for the MM's blind reception and detection based on the MM EAC.

Finally, we present analytical bounds on the BLER performance of the MM scheme over an AWGN channel when using FEC with linear block coding.

Before proceeding to the analytical study of MM we will tackle the subject of multirate processing of random signals, namely of *stationary* and *cyclostationary* signals [93, 107]. The acquired knowledge will be particularly useful in the analysis that follows.

This chapter is partly based on [38] and [39].

5.1 MULTIRATE PROCESSING OF RANDOM SIGNALS

The statistical analysis of random signals and noise is traditionally carried out assuming *wide sense stationarity* (WSS). In multirate digital signal processing, we often encounter time-varying linear systems such as decimators and modulators. In many applications these building blocks are interconnected with linear filters to form more complex systems. It is often necessary to understand how the statistical behavior of a discrete-time signal changes as it passes through such systems.

5.1.1 *Stationarity Definitions*

When analyzing a digital system, input and output signals are usually random and therefore modeled as discrete-time *stochastic random processes* [45, 84, 87]. A discrete-time random process $x[n]$ is simply an indexed sequence of random variables characterized statistically by their joint pdf, with

$$p(x[n]\,; n = n_1, n_2, \cdots, n_k) \equiv p\big(x_{n_1}, x_{n_2}, \cdots, x_{n_k}\big)\ , \tag{5.1}$$

denoting k-order joint *pdf*, for the k random variables obtained from observing $x[n]$ at a given set of time instants[1] $\{n_1, n_2, \cdots, n_k\}$.

A stochastic process (continuous- or discrete-time) is called *strict sense stationary* (SSS) if its statistical properties are invariant over time. From the definition, $x[n]$ is said to be an SSS stochastic process if and only if,

$$p\big(x_{n_1}, x_{n_2}, \cdots, x_{n_k}\big) = p\big(x_{n_1+c}, x_{n_2+c}, \cdots, x_{n_k+c}\big)\ , \tag{5.2}$$

for all $k \in \mathbb{N}$ and $c \in \mathbb{Z}$.

However, a less stringent condition of stationarity is often used in practice. A stochastic process $x[n]$ is called *wide sense stationary* (WSS) if its mean is constant

$$E\{x[n]\} = \eta\ , \tag{5.3}$$

and if its autocorrelation, defined as

$$r_x[n;\tau] = E\{x[n+\tau]\, x^*[n]\}\ , \tag{5.4}$$

is only a function of the time *lag* between random variables $x[n+\tau]$ and $x[n]$, and it is also independent of the specified time index n, i.e.

$$r_x[n;\tau] = r_x[n+c;\tau] = r_x[\tau] \quad , \forall\, c \in \mathbb{Z}\ . \tag{5.5}$$

Two stochastic processes $x[n]$ and $y[n]$ are said to be *jointly* WSS if $x[n]$ and $y[n]$ are WSS and the cross-correlation

$$r_{xy}[n;\tau] = E\{x[n+\tau]\, y^*[n]\}\ , \tag{5.6}$$

depends only on the time lag τ, and so

$$r_{xy}[n;\tau] = r_{xy}[n+c;\tau] = r_{xy}[\tau]\,, \quad \text{for any } \ c \in \mathbb{Z}\ . \tag{5.7}$$

1 The random variable obtained from the evaluation of $x[n]$ at a given time instant n_i, is denoted as $x[n_i] \equiv x_{n_i}$

Another important definition in stationarity, particularly useful in the analysis of multirate systems, is the concept of *cyclostationarity* [22, 25, 84, 93]. A stochastic process $x[n]$ is said to be *wide sense cyclostationary* (WSCS) with period $L \in \mathbb{N}$ (denoted as $\text{WSCS}_{(L)}$) if

$$E\{x[n]\} = E\{x[n+kL]\} \quad , \forall k \in \mathbb{Z} \, , \tag{5.8}$$

and

$$r_x[n;\tau] = r_x[n+kL;\tau] \quad , \forall k \in \mathbb{Z} \, , \tag{5.9}$$

i.e. the mean and autocorrelation of $x[n]$ are both periodic functions in time with period L.

5.1.2 *Average Power of WSS and WSCS stochastic processes*

The *instantaneous power* of a zero mean stochastic process $x[n]$ is defined as[2]

$$\sigma_x^2[n] = E\{x[n]\, x^*[n]\} = r_x[n;0] \ . \tag{5.10}$$

If $x[n]$ is WSS, the autocorrelation $r_x[n;0]$ depends only on the time lag $\tau = 0$ in accordance with (5.5), and so its power is constant over time. Hence, the average power of a WSS stochastic process $x[n]$ is given by

$$\sigma_x^2 = r_x[0] \ . \tag{5.11}$$

However, if $x[n]$ is $\text{WSCS}_{(L)}$ its instantaneous power varies periodically over time with period L. In order to compute the average power of a WSCS process we must map the process to an equivalent stationary process. Since autocorrelation $r_x[n;\tau]$ is a periodic function of time it can be represented by a Fourier series [22, 25, 93, 107]

$$r_x[n;\tau] = \sum_{k=0}^{L-1} R_x[k;\tau]\, e^{jk\frac{2\pi}{L}n} \, , \tag{5.12}$$

with Fourier coefficients, called *cyclic autocorrelations*, given by

$$R_x[k;\tau] = \frac{1}{L} \sum_{n=0}^{L-1} r_x[n;\tau]\, e^{-jk\frac{2\pi}{L}n} \ . \tag{5.13}$$

2 $\sigma_x^2[n]$ usually defines the variance of a r.v., i.e. $\sigma_x^2[n] = E\left\{|x[n] - E\{x[n]\}|^2\right\}$. For zero mean random processes the *autocovariance* and autocorrelation are equal, and so (5.11) is verified [45]. For convenience, unless stated otherwise, random processes are assumed to have zero mean.

Sequence $x[n]$ can be mapped to an equivalent stationary process by computing the time average autocorrelation function, which is simply the cyclic autocorrelation $R_x[k;\tau]$ at $k = 0$. The stationarization is equivalent to filtering the $WSCS_{(L)}$ process with an ideal low-pass filter with cut-off frequency π/L [93, 107]. The autocorrelation of the "stationary component" of $x[n]$ is thus given by

$$\bar{r}_x[\tau] = R_x[0;\tau] = \frac{1}{L}\sum_{n=0}^{L-1} r_x[n;\tau] \,, \tag{5.14}$$

and the average power of the WSCS process $x[n]$ is calculated as

$$\sigma_x^2 = \bar{r}_x\,[\tau]|_{\tau=0} = R_x[0;\tau]|_{\tau=0} = \frac{1}{L}\sum_{n=0}^{L-1} r_x[n;0] \,. \tag{5.15}$$

5.1.3 *Stationarity under Rate Change*

In this section we present some of the propositions that characterize the changes in the statistical behavior of a stochastic signal due to a sampling rate change. The subject is dealt with in detail in [93]. The presented propositions will be particularly useful in the analytical analysis of MM systems.

Proposition 5.1. *If we pass a WSS signal* $x[n]$ *through an* L*-fold up-sampler (see figure 19.(a)), the output* $y_u[n]$ *is* $WSCS_{(L)}$.

Proof. We know that the up-sampled signal $y_u[n]$ relates to $x[n]$ according to (4.1), which can also be written as

$$y[n] = \sum_{k=-\infty}^{\infty} x[k]\,\delta[n - kL] \,. \tag{5.16}$$

Making use of (5.16), we compute the autocorrelation of $y_u[n]$

$$\begin{aligned} r_{y_u}[n;\tau] &= E\{y_u[n+\tau]\,y_u^*[n]\} \\ &= E\left\{\sum_i \sum_j x[i]\,x^*[j]\,\delta[n+\tau-iL]\,\delta[n-jL]\right\} \,. \end{aligned} \tag{5.17}$$

Exchanging the expectation operator wit both sums[3], and since $x[n]$ is WSS, it results that

3 Expectation, $E\{\cdot\}$, is a linear operation.

$$r_{y_u}[n;\tau] = \sum_i \sum_j E\{x[i]\,x^*[j]\}\,\delta[n+\tau-iL]\,\delta[n-jL]$$
$$= \sum_i \sum_j r_x[i-j]\,\delta[n+\tau-iL]\,\delta[n-jL]\ . \tag{5.18}$$

It is straightforward to conclude from the analysis of (5.18) that

$$r_{y_u}[n;\tau] = r_{y_u}[n+kL;\tau] \quad ,\ \forall k \in \mathbb{Z} \tag{5.19}$$

i.e. it is periodic in time with period L and, therefore, $y_u[n]$ is $\text{WSCS}_{(L)}$ as we wanted to prove.

□

Based on proposition 5.1 we can easily relate the average power of WSCS signal $y_u[n]$ to the average power of the up-sampler input $x[n]$. By using (5.15), we can compute $\sigma^2_{y_u}$ as

$$\sigma^2_{y_u} = \frac{1}{L}\sum_{n=0}^{L-1} r_{y_u}[n;0]\ . \tag{5.20}$$

According to (5.18) we have

$$r_{y_u}[n;0] = \begin{cases} r_x[0] & ,\ n = kL \\ 0 & ,\ n \neq kL \end{cases} \quad ,\ \forall k \in \mathbb{Z}\ , \tag{5.21}$$

which gives

$$\sigma^2_{y_u} = \frac{r_x[0]}{L} = \frac{\sigma^2_x}{L}\ . \tag{5.22}$$

Proposition 5.2. *If we pass a $\text{WSCS}_{(P)}$ signal $x[n]$ through an L-fold up-sampler (see figure 19.(a)), the output $y_u[n]$ is $\text{WSCS}_{(LP)}$, and the power of the output signal relates to the power at up-sampler entrance as $\sigma^2_{y_u} = \sigma^2_x/L$.*

Proof. Given that $x[n]$ is now $\text{WSCS}_{(P)}$, (5.17) can be simplified as

$$r_{y_u}[n;\tau] = \sum_i \sum_j r_x[j;i-j]\,\delta[n+\tau-iL]\,\delta[n-jL]\ . \tag{5.23}$$

In order to prove that $y_u[n]$ is WSCS$_{(LP)}$, we simply apply the definition (5.9) of a WSCS process and compute $r_{y_u}[n+kLP;\tau]$, for any $k \in \mathbb{Z}$. Hence, according to (5.23) we get

$$r_{y_u}[n+kLP;\tau] = \sum_i \sum_j r_x[j;i-j]\,\delta[n+\tau-(i-kP)L]\,\delta[n-(j-kP)L]\ , \quad (5.24)$$

and by performing a change of variables, $i' = i-kP$ and $j' = j-kP$, it results

$$r_{y_u}[n+kLP;\tau] = \sum_{i'} \sum_{j'} r_x\left[j'+kP;i'-j'\right]\,\delta\left[n+\tau-i'L\right]\,\delta\left[n-j'L\right]\ . \quad (5.25)$$

Since $x[n]$ is WSCS$_{(P)}$ (5.25) simplifies to

$$\begin{aligned} r_{y_u}[n+kLP;\tau] &= \sum_{i'} \sum_{j'} r_x\left[j';i'-j'\right]\,\delta\left[n+\tau-i'L\right]\,\delta\left[n-j'L\right] \\ &= r_{y_u}[n;\tau]\ , \end{aligned} \quad (5.26)$$

proving that $y_u[n]$ is WSCS$_{(LP)}$.

The power of the output signal can be computed according to (5.15) and so considering (5.23) σ_y^2 is given by

$$\begin{aligned} \sigma_{y_u}^2 &= \frac{1}{LP}\sum_{n=0}^{LP-1} r_{y_u}[n;0] \\ &= \frac{1}{LP}\sum_{n=0}^{LP-1}\sum_{i=-\infty}^{\infty}\sum_{j=-\infty}^{\infty} r_x[j;i-j]\,\delta[n-iL]\,\delta[n-jL]\ . \end{aligned} \quad (5.27)$$

The product $\delta[n-iL]\times\delta[n-jL]$ is zero for all $i \neq j$ and so (5.27) simplifies to

$$\begin{aligned} \sigma_{y_u}^2 &= \frac{1}{LP}\sum_{n=0}^{LP-1}\sum_{j=-\infty}^{\infty} r_x[j;0]\,\delta[n-jL] \\ &= \frac{1}{LP}\sum_{j=-\infty}^{\infty}\left(r_x[j;0]\sum_{n=0}^{LP-1}\delta[n-jL]\right)\ , \end{aligned} \quad (5.28)$$

where, $\sum_{n=0}^{LP-1}\delta[n-jL] = 1$ for $j = 0,\cdots,P-1$, and is zero for the remaining values of j. Relation between $\sigma_{y_u}^2$ and σ_x^2 thus follows directly

$$\sigma_{y_u}^2 = \frac{1}{LP}\sum_{j=0}^{P-1} r_x[j;0] = \frac{\sigma_x^2}{L}\ . \quad (5.29)$$

□

Some considerations on propositions 5.1 and 5.2 and their proofs should be noted. Firstly, proposition 5.1 and result (5.22) are a particular case of proposition 5.2, in view of the fact that a WSS stochastic process can be interpreted as being $\text{WSCS}_{(P)}$ with $P = 1$. Secondly, when proving the WSCS property of the up-sampled signal $y_u[n]$ we have ignored condition (5.8); given that $E\{y_u[n]\}$ is periodic, the proof is trivial.

Proposition 5.3. *If we pass a* $\text{WSCS}_{(L)}$ *signal* $x[n]$ *through an* M*-fold down-sampler (see figure 19.(b)), the output* $y_d[n]$ *is* $\text{WSCS}_{(K)}$*, where* $K = L/\gcd(L, M)$.

Proof. Using the input-output relationship of a down-sampler (4.5) we can write the autocorrelation of $y_d[n]$ as

$$\begin{aligned} r_{y_d}[n;\tau] &= E\{y_d[n+\tau]\, y_d^*[n]\} \\ &= E\{x[nM+\tau M]\, x^*[nM]\} = r_x[nM;\tau M] \ . \end{aligned} \tag{5.30}$$

Assuming that $y_d[n]$ is $\text{WSCS}_{(K)}$ from (5.9) we know that

$$r_{y_d}[n;\tau] = r_{y_d}[n+K;\tau] \ , \tag{5.31}$$

and so by substituting (5.30) in (5.31), it results

$$r_x[nM;\tau M] = r_x[nM+KM;\tau M] \ . \tag{5.32}$$

Since $x[n]$ is $\text{WSCS}_{(L)}$ (5.32) is satisfied if

$$KM = iL \ , \tag{5.33}$$

for some $i \in \mathbb{Z}$. The smallest value of $K \in \mathbb{N}$ for which there is an integer i that satisfies (5.33) is given by

$$K = \frac{L}{\gcd(L, M)} \ , \tag{5.34}$$

as we wanted to prove.

□

Note that a WSS stochastic process is a special case of a $\text{WSCS}_{(L)}$ with $L = 1$. Hence, according to proposition 5.3, the down-sampling of a WSS process always gives rise to a WSS signal.

5.1.4 *Filtering Random Processes*

Linear time-invariant (LTI) filters [76, 82] are frequently used to perform a variety of different signal processing tasks. It is therefore important to be able to determine how the statistics of a random process change as a result of being filtered by an LTI system.

A digital LTI filter is usually characterized by its impulse response $h[n]$ or, equivalently, by its frequency response $H(z)$. The output $y[n]$ of the system may be expressed as the convolution sum of the input $x[n]$ with $h[n]$, i.e.

$$y[n] = x[n] * h[n] = \sum_{k=-\infty}^{\infty} h[k]\, x[n-k] \ . \tag{5.35}$$

Proposition 5.4. *The output response of an LTI digital system to a WSCS$_{(L)}$ signal, $x[n]$, is a WSCS$_{(L)}$ process $y[n]$ with autocorrelation*

$$r_y[n;\tau] = \sum_{i=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} h[i]\, h^*[j]\, r_x[n-j;\, i-j-\tau] \ , \tag{5.36}$$

where $h[n]$ is the LTI system's impulse response.

Proof. Let's begin by computing the mean value of $y[n]$ which is given by

$$E\{y[n]\} = E\left\{ \sum_{k=-\infty}^{\infty} h[k]\, x[n-k] \right\} = \sum_{k=-\infty}^{\infty} h[k]\, E\{x[n-k]\} \ . \tag{5.37}$$

Signal $x[n]$ is WSCS$_{(L)}$, and so $E\{x[n]\}$ is periodic with period L. Hence, it results from (5.37) that $E\{y[n]\}$ is also periodic with the same period. In fact, by computing $E\{y[n+iL]\}$ for any $i \in \mathbb{Z}$ we conclude that

$$\begin{aligned} E\{y[n+iL]\} &= \sum_{k=-\infty}^{\infty} h[k]\, E\{x[n-k+iL]\} \\ &= \sum_{k=-\infty}^{\infty} h[k]\, E\{x[n-k]\} = E\{y[n]\} \ . \end{aligned} \tag{5.38}$$

The autocorrelation function of the output $y[n]$ is

$$
\begin{aligned}
r_y[n;\tau] &= E\{y[n+\tau]\, y^*[n]\} \\
&= E\left\{ \sum_{i=-\infty}^{\infty} h[i]\, x[n+\tau-i] \times \sum_{j=-\infty}^{\infty} h^*[j]\, x^*[n-j] \right\} \\
&= \sum_{i=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} h[i]\, h^*[j]\, r_x[n-j; j-i+\tau] \;, \qquad (5.39)
\end{aligned}
$$

which is also a periodic function in n. By evaluating $r_y[n+kL]$ for any $k \in \mathbb{Z}$, and making use of the $\text{WSCS}_{(L)}$ property of $x[n]$, it results that

$$
\begin{aligned}
r_y[n+kL;\tau] &= \sum_{i=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} h[i]\, h^*[j]\, r_x[n-j+kL; j-i+\tau] \\
&= \sum_{i=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} h[i]\, h^*[j]\, r_x[n-j; j-i+\tau] = r_y[n;\tau] \;. \qquad (5.40)
\end{aligned}
$$

Given (5.38) and (5.40), we conclude, according to the WSCS definition, that $y[n]$ is $\text{WSCS}_{(L)}$.

□

For the particular instance of the LTI system's input being WSS (i.e. $\text{WSCS}_{(1)}$) then output $y[n]$ is also WSS and its autocorrelation (5.36) simplifies to

$$
r_y[\tau] = \sum_{i=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} h[i]\, h^*[j]\, r_x[j-i+\tau] = h[\tau] * h^*[-\tau] * r_x[\tau] \;, \qquad (5.41)
$$

with $h[\tau] * h^*[-\tau]$ being called the *filter aperiodic autocorrelation* [83].

5.1.5 *The Interpolator Case*

So far we have seen the effects of basic multirate building blocks on stationary random inputs. But it is interesting to study similar properties for some standard interconnections of these building blocks. We will only examine the interpolator filter structure (studied in section 4.1), since it is part of any typical SC transmitter (figures 1 and 10).

Some interesting properties and theorems on the interpolator structure have been stated in [22, 93, 107]. Recalling figure 24 (see chapter 4), assume that $x[n]$ is WSS. Hence, according to propositions 5.1 and 5.4, the up-sampled signal $x_u[n]$ and the interpolator's output $y[n]$ are $\text{WSCS}_{(L)}$. Nevertheless, Sathe *et al.* [93] have proved that if the LTI filter $H(z)$ is such that the L-fold decimation of

its impulse response does not create aliasing then the interpolated signal $y[n]$ is WSS.

In the generic block diagram of a typical SC transmitter presented in figure 10 (see chapter 3), the pulse shaping filter $H(z)$ (part of an interpolator structure), does not necessarily have a spectrum shape with single side bandwidth limited to π/L, and so its output may not be WSS. However, $H(z)$ should satisfy the Nyquist condition for zero ISI with pulse duration of L samples (interpolator up-sampling factor) [44, 65, 87]. This means that the aperiodic autocorrelation of the filter

$$h_c[n] = h[n] * h^*[-n] \,, \tag{5.42}$$

is such that,

$$h_c[kL] = \begin{cases} \sum_{n=-\infty}^{\infty} |h[n]|^2 & , \, k = 0 \\ 0 & , \, k \neq 0 \end{cases} \quad , \, \forall \, k \in \mathbb{Z} \,. \tag{5.43}$$

To this special case of a Nyquist interpolator filter we can relate the average power of a WSS or WSCS signal $x[n]$ fed into the interpolator to the average power of the WSCS output signal $y[n]$.

Theorem 5.1. *Consider a multirate interpolator system formed by an* L*-fold up-sampler followed by a pulse shaping filter with impulse response* $h[n]$*, whose aperiodic autocorrelation satisfies the Nyquist pulse criterion for zero ISI with pulse duration of* L *samples. The interpolated signal's average output power,* σ_y^2*, in response to a zero mean WSCS*$_{(P)}$ *stochastic process* $x[n]$ *with power* σ_x^2*, is given by*

$$\sigma_y^2 = \frac{\sigma_x^2}{L} \sum_{n=-\infty}^{\infty} |h[n]|^2 \,. \tag{5.44}$$

Proof. Assuming that input signal $x[n]$ is WSCS$_{(P)}$, it follows from propositions 5.2 and 5.4 that the L-fold up-sampled signal $x_u[n]$ and the output pulse shaped signal $y[n]$ are both WSCS$_{(LP)}$ processes.

The pulse shaped signal relates to input as

$$\begin{aligned} y[n] = h[n] * x_u[n] &= h[n] * \sum_{k=-\infty}^{\infty} x[k] \, \delta[n - kL] \\ &= \sum_{k=-\infty}^{\infty} x[k] \, h[n - kL] \,, \end{aligned} \tag{5.45}$$

whereby, computing $r_y[n;\tau]$ according with definition (5.4), we obtain

$$r_y[n;\tau] = \sum_{i=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} r_x[j;i-j]\, h[n+\tau-iL]\, h^*[n-jL] \,. \tag{5.46}$$

The pulse shaped signal is a $\text{WSCS}_{(LP)}$ process and so its power is computed as

$$\begin{aligned} \sigma_y^2 &= \frac{1}{LP} \sum_{n=0}^{LP-1} r_y[n;0] \\ &= \frac{1}{LP} \sum_{n=0}^{LP-1} \sum_{i=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} r_x[j;i-j]\, h[n-iL]\, h^*[n-jL] \,, \end{aligned} \tag{5.47}$$

whereby, performing a variable change, $i = j + a$, it results

$$\sigma_y^2 = \frac{1}{LP} \sum_{n=0}^{LP-1} \sum_{a=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} r_x[j;a]\, h[n-(j+a)\,L]\, h^*[n-jL] \,. \tag{5.48}$$

Input $x[n]$ is $\text{WSCS}_{(P)}$. Hence, $r_x[j;a] = r_x[j+mP;a]$, for all $m \in \mathbb{Z}$ and (5.48) can thus be rewritten as

$$\begin{aligned} \sigma_y^2 &= \frac{1}{LP} \sum_{n=0}^{LP-1} \sum_{a=-\infty}^{\infty} \sum_{j=0}^{P-1} \left(r_x[j;a] \sum_{m=-\infty}^{\infty} h[n-(j+mP+a)L]\, h^*[n-(j+mP)L] \right) \\ &= \frac{1}{LP} \sum_{a=-\infty}^{\infty} \sum_{j=0}^{P-1} \left(r_x[j;a] \sum_{n=0}^{LP-1} \sum_{m=-\infty}^{\infty} h[n-mPL-(j+a)L]\, h^*[n-mPL-jL] \right) \\ &= \frac{1}{LP} \sum_{a=-\infty}^{\infty} \sum_{j=0}^{P-1} \left(r_x[j;a] \sum_{n'=-\infty}^{\infty} h\big[n'-jL-aL\big]\, h^*\big[n'-jL\big] \right) \quad \Leftarrow \begin{matrix}\text{variable change} \\ k=n'-jL\end{matrix} \\ &= \frac{1}{LP} \sum_{a=-\infty}^{\infty} \sum_{j=0}^{P-1} \left(r_x[j;a] \sum_{k=-\infty}^{\infty} h[k-aL]\, h^*[k] \right) , \end{aligned} \tag{5.49}$$

where the sum inside the parenthesis is simply the filter aperiodic autocorrelation (5.42) at $n=-aL$, where we get

$$\sigma_y^2 = \frac{1}{L} \sum_{a=-\infty}^{\infty} \left(\frac{1}{P} \sum_{j=0}^{P-1} r_x[j;a]\, h_c\,[-aL] \right) . \tag{5.50}$$

Filter $h[n]$ is assumed to satisfy the Nyquist pulse criterion for zero ISI, and so, accordingly to (5.43), $h_c[-aL]$ it is only nonzero for $a = 0$ wherefore, theorem statement results straightforward:

$$\sigma_y^2 = \frac{1}{L}\left(\frac{1}{P}\sum_{j=0}^{P-1} r_x[j;0]\right) h_c[0] = \frac{\sigma_x^2}{L}\sum_{n=-\infty}^{\infty} |h[n]|^2 \,. \tag{5.51}$$

□

5.2 ANALYTICAL POWER ANALYSIS

In this section we analyze the statistical properties of the MM's factors sequence $m[n]$ and of the average power of the MM signal. Although attention will mainly focus on the MPMM scheme (figure 29), given its excellent performance, the analysis is also valid for the LUT-MM approach. Here the statistical distributions of the MM factors can be known exactly from the information stored in the LUT.

Let S_i, with $i = 1, \cdots, M$, denote the M symbols of an M-ary constellation $\mathcal{X}$, and refer back to figure 29 in section 4.2. The sequence of symbols[4] $s[n]$ at the output of the modulator is assumed to be a discrete-time zero mean WSS random process. The symbols are considered to be independent and identically distributed (i.i.d.) random variables (r.v.) with average transmitted power given by

$$\sigma_s^2 = E\left\{|s[n]|^2\right\} = \frac{1}{M}\sum_{i=1}^{M} |S_i|^2 = \varepsilon_s \,, \tag{5.52}$$

i.e. it equals the average energy of a constellation symbol, ε_s. It is also assumed that MM sequence, $s_{MM}[n] = m[n] \times s[n]$, preserves the WSS properties of $s[n]$.

5.2.1 *Statistical Distributions of MM factors*

Although the computation of m_n depends on s_n and its 2N closer neighbors (as seen in chapters 3 and 4), it is also true that s_n is the symbol that contributes most to the evaluation of m_n. For reasons of treatability of the problem we will assume that the pdf of r.v. m_n only depends on the current constellation symbol s_n being MM, i.e.

$$p(m_n \mid s_{n-N}, \cdots, s_n, \cdots, s_{n+N}) \simeq p(m_n \mid s_n) \,. \tag{5.53}$$

4 In order to avoid notation misinterpretation, we recall that s_n denotes a transmitted constellation symbol at a given time instant n, whereas when a capital letter is used, e.g. S_i, we are referring to the i^{th} symbol of the M-ary constellation. (See also footnote 2 on page 52.)

Even with this simplification, because of the recursive computation nature of each MM factor that depends on previous computed MM factors, and the nonlinearity of equations (4.34), (4.44) and (4.45), it is not possible to arrive at a close expression for the *pdf* of m_n by analytical procedures.

By performing a goodness of fit statistical analysis, we conclude that the conditional *pdf* of m_n, given that s_n is constellation symbol S_i, can be statically handled using a beta distribution [84] (as expected, since $m_n \in]0, 1]$), and so it can be expressed[5] as

$$p(m_n \mid s_n = S_i) = p(m \mid S_i) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\,\Gamma(\beta)} m^{\alpha-1}\,(1-m)^{\beta-1} \quad , \tag{5.54}$$

where

$$\Gamma(\alpha) = \int_0^\infty x^{\alpha-1} e^{-x} dx \, , \tag{5.55}$$

represents the gamma function. The beta distribution model also matches the LUT-MM case since $0 < m \leqslant 1$.

Due to constellation and RRC symmetry properties, the pdf of m is only conditioned in $|S_i|$ and so (5.54) is the same for all the symbols in $\mathcal{X}$ having the same magnitude, as shown in figure 41 for 16-QAM. Therefore, let $\mathbf{S}^{A_j}$ be the sub-set of symbols of the constellation $\mathcal{X}$ with amplitude A_j, and $n_{A_j} = \#\left\{\mathbf{S}^{A_j}\right\}$. Then for $S_i \in \mathbf{S}^{A_j}$, (5.54) is hereafter written as

$$p\left(m \mid S_i \in \mathbf{S}^{A_j}\right) = \frac{\Gamma\left(\alpha_{A_j} + \beta_{A_j}\right)}{\Gamma\left(\alpha_{A_j}\right)\Gamma\left(\beta_{A_j}\right)} m^{\alpha_{A_j}-1}\,(1-m)^{\beta_{A_j}-1} \quad , \tag{5.56}$$

where parameters α_{A_j} and β_{A_j} are a function of the constellation geometry and the RRC roll-off. The dependence on the number of MPMM stages, k, vanishes for $k \geqslant 2$ since the need to back-off is almost completely eliminated, as has been shown in section 4.3. The same conclusion on the MPMM's efficiency can be drawn from the observation of the complementary cumulative distribution functions (CCDF) of the amplitude of the transmitted MM signal, presented in figure 42 for the demanding case of 0.1 roll-off RRC filtering. Typical values for α_{A_j} and β_{A_j} were thus obtained through *maximum likelihood estimation* [71], with the beta distribution as prototype, considering a large data set of MM values obtained through simulation for different constellations and RRC's roll-offs with a 3-stage MPMM configuration. These parameters are presented in table 5.

5 Temporal subscript n can thus be dropped, since under assumption (5.53) the statistical distribution that m_n follows, depends only on the constellation symbol S_i and not on the particular instant n.

		Roll-off →	0.1		0.2		0.3	
		n_{A_j}	α_{A_j}	β_{A_j}	α_{A_j}	β_{A_j}	α_{A_j}	β_{A_j}
BPSK	$A_1 = 1$	2	6.8	3.09	9.9	3.99	13.0	4.64
8-PSK	$A_1 = 1$	8	9.8	3.15	14.8	4.21	21.2	5.39
16-APSK	$A_1 = 0.36056$	4	18.5	1.49	27.4	1.95	37.2	2.50
	$A_2 = 1.13578$	12	8.6	2.09	12.4	2.76	17.9	3.69
16-QAM	$A_1 = 0.44721$	4	14.8	0.51	18.2	0.53	19.7	0.52
	$A_2 = 1$	8	7.3	0.48	9.9	0.49	13.3	0.50
	$A_3 = 1.34164$	4	7.9	1.38	10.9	1.81	14.4	2.30
32-APSK	$A_1 = 0.24227$	4	18.6	0.89	24.5	1.08	29.2	1.27
	$A_2 = 0.68807$	12	13.7	0.80	19.4	0.99	24.7	1.17
	$A_3 = 1.2768$	16	8.2	1.58	11.9	2.15	16.9	2.92
64-QAM	$A_1 = 0.21821$	4	14.8	0.19	18.1	0.18	20.5	0.18
	$A_2 = 0.48795$	8	13.8	0.19	17.4	0.18	20.0	0.18
	$A_3 = 0.65465$	4	12.4	0.19	16.0	0.18	18.6	0.18
	$A_4 = 0.78679$	8	10.8	0.19	14.9	0.18	18.2	0.18
	$A_5 = 0.89973$	8	9.1	0.19	12.7	0.18	16.6	0.18
	$A_6 = 1.09108$	12	6.3	0.19	9.0	0.18	13.4	0.18
	$A_7 = 1.17513$	8	5.3	0.20	7.2	0.18	10.6	0.18
	$A_8 = 1.32736$	8	4.0	0.23	4.9	0.22	6.3	0.23
	$A_9 = 1.52752$	4	6.6	0.88	8.8	1.17	11.8	1.60

Table 5: Parameters of beta pdf that models MM values' distributions. The set of amplitudes A_j provided for each constellation, assumes that $\varepsilon_s = 1$.

5.2.2 *Average Power and PAPR Reduction*

For a better understanding, and without loss of generality, we will consider that the multirate pulse shaping system (L up-sampler followed by the RRC filter $h[n]$) has unitary power gain, i.e. $G_{RRC} = 1$, and so according to theorem 5.1

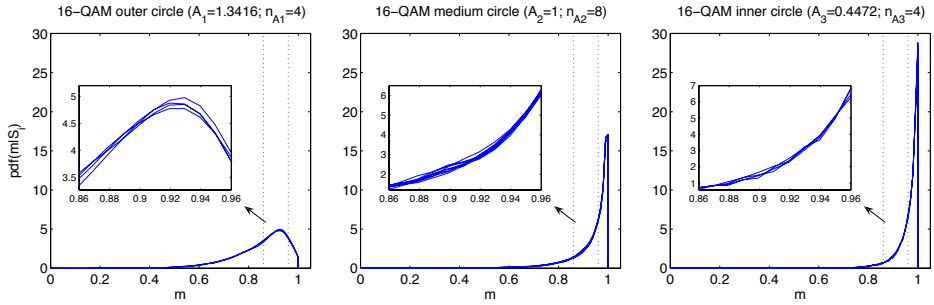


Figure 41: Conditional pdf of MM coefficients given that constellation symbol S_i is being MM for the 16-QAM case. Distributions for equal magnitude symbols are plotted in the same graph.

and propositions 5.1 and 5.4, the bandwidth limited MM signal $x[n]$ is $WSCS_{(L)}$ and has power

$$\sigma_x^2 = \sigma_{s_{MM}}^2 = \sum_{i=1}^{M} \left(|S_i|^2 \Pr(S_i) \int_0^1 m^2 p(m \mid S_i)\, dm \right) . \tag{5.57}$$

Considering (5.56) the power of the transmitted signal for a M-ary constellation can thus be expressed as

$$\begin{aligned} \sigma_{s_{MM}}^2 &= \sum_j \left(\frac{n_{A_j} A_j^2}{M} \int_0^1 m^2 p\left(m \mid \mathbf{S}^{A_j}\right) dm \right) \\ &= \sum_j \left(\frac{n_{A_j} A_j^2}{M} \int_0^1 \frac{\Gamma\left(\alpha_{A_j} + \beta_{A_j}\right)}{\Gamma\left(\alpha_{A_j}\right)\Gamma\left(\beta_{A_j}\right)} m^{\alpha_{A_j}+1} (1-m)^{\beta_{A_j}-1}\, dm \right) , \end{aligned} \tag{5.58}$$

and from the integral calculation,

$$\sigma_{s_{MM}}^2 = \sum_j \left(\frac{n_{A_j} A_j^2}{M} \times \frac{\left(\alpha_{A_j}+1\right)\alpha_{A_j}}{\left(\alpha_{A_j}+\beta_{A_j}+1\right)\left(\alpha_{A_j}+\beta_{A_j}\right)} \right) , \tag{5.59}$$

which, for constant amplitude constellations, can be further simplified as

$$\sigma_{s_{MM}}^2 = \frac{(\alpha+1)\,\alpha}{(\alpha+\beta+1)\,(\alpha+\beta)} \varepsilon_s . \tag{5.60}$$

	Analytical Model			Simulation		
Roll-off →	0.1	0.2	0.3	0.1	0.2	0.3
8-PSK	0.58588	0.61474	0.64153	0.59080	0.61670	0.64149
16-APSK	0.66713	0.68459	0.69970	0.66917	0.68522	0.69867
16-QAM	0.82198	0.83817	0.85073	0.82175	0.83741	0.84875
32-APSK	0.74904	0.75960	0.76642	0.75093	0.76034	0.76571
64-QAM	0.91685	0.93032	0.93827	0.91529	0.92915	0.93711

Table 6: Average power of the bandwidth limited MM transmitted signal for a 3-stage MPMM configuration, considering different normalized constellations, i.e. $\varepsilon_s = 1$, and a multirate pulse shaping filter with $G_{RRC} = 1$.

Table 6 shows that (5.59), computed for α_{A_j} and β_{A_j} provided in table 5, accurately follows the simulation results. The results presented in table 6 relate to normalized constellations, i.e. $\varepsilon_s = 1$, and $G_{RRC} = 1$.

The *average power loss* due to MM is thus given by:

$$\begin{aligned}\overline{P}_{MM\_loss} &= 10\log_{10}\left(\sigma^2_{s_{MM}}/\sigma^2_s\right) \\ &= 10\log_{10}\sum_j\left(\frac{n_{A_j}A_j^2}{M\varepsilon_s}\times\frac{\left(\alpha_{A_j}+1\right)\alpha_{A_j}}{\left(\alpha_{A_j}+\beta_{A_j}+1\right)\left(\alpha_{A_j}+\beta_{A_j}\right)}\right)\quad(\text{dB})\,,\end{aligned} \tag{5.61}$$

where we have made use of (5.52) and (5.59).

Although this power loss gives rise to some increased sensitivity to noise, we must consider the power efficiency gain that results from eliminating the need to back-off from HPA saturation. Overall *power efficiency gain for MM* system, G_{MM}, can be evaluated as the observed reduction in PAPR, i.e.

$$G_{MM} = PAPR_{no\_MM} - PAPR_{with\_MM}\,. \tag{5.62}$$

As stated in chapter 1, it is assumed that the HPA should be able to handle at least $PAPR_{Const}$ and, therefore, the MM polar threshold used in simulations and with respect to which the required back-off is measured is set to be

$$A^2 = \max_{S_i\in\mathcal{X}}|S_i|^2\times G_{RRC}\,, \tag{5.63}$$

and so, using (5.63) in the manipulation of (5.62) we get[6]

6 Note that we are assuming $G_{RRC} = 1$, without any loss of generalization.

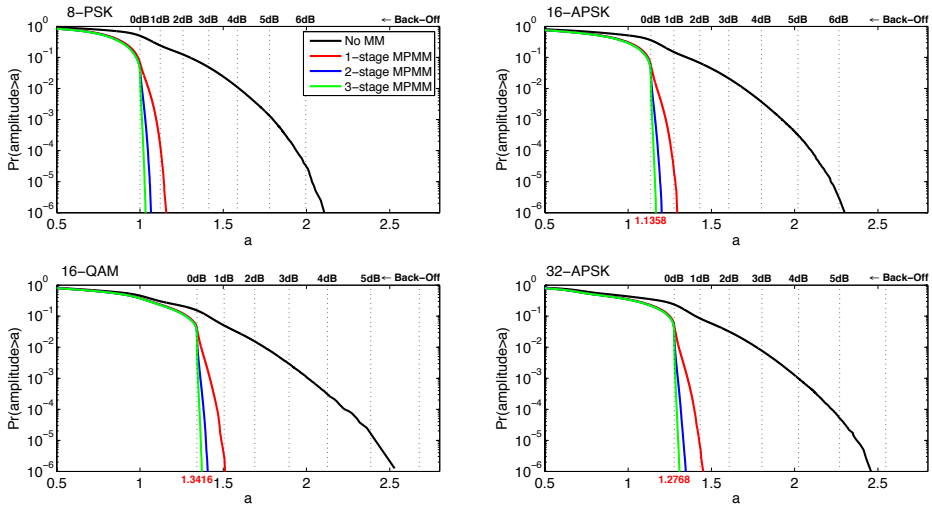


Figure 42: CCDF of the transmitted signal amplitude, $|x[n]|$, and required back-off, considering MPMM with a different number of stages for some constellations in the demanding case of 0.1 roll-off RRC filtering.

$$
\begin{aligned}
G_{MM} &= PAPR_{no\_MM} - 10\log_{10}\left(\frac{A^2}{\sigma^2_{s_{MM}}}\right) \qquad \Leftarrow \quad G_{RRC}=1 \\
&= PAPR_{no\_MM} - 10\log_{10}\left(\frac{\max\limits_{S_i\in\mathcal{X}}|S_i|^2}{\sigma_s^2}\right) - 10\log_{10}\left(\frac{\sigma_s^2}{\sigma^2_{s_{MM}}}\right) \\
&= PAPR_{no\_MM} - PAPR_{Const} - \overline{P}_{MM\_loss} \; . \qquad (5.64)
\end{aligned}
$$

Bearing in mind that, as stated in section 1.2 and proved in appendix A,

$$
PAPR_{no\_MM} = PAPR_{Const} + PAPR_{RRC} \; , \qquad (5.65)
$$

by substituting (1.4) and (5.61) in equation (5.64), we get

$$
\begin{aligned}
G_{MM} &= PAPR_{RRC} - \overline{P}_{MM\_loss} \\
&= 10\log_{10}\left[\frac{\max\limits_{i=0,\cdots,L-1}\left(\sum\limits_n |h[nL+i]|\right)^2}{\sum\limits_j\left(\frac{n_{A_j}A_j^2}{M\varepsilon_s}\times\frac{\left(\alpha_{A_j}+1\right)\alpha_{A_j}}{\left(\alpha_{A_j}+\beta_{A_j}+1\right)\left(\alpha_{A_j}+\beta_{A_j}\right)}\right)}\right] \quad (\text{dB}) \; . \qquad (5.66)
\end{aligned}
$$

Expression (5.66) is indeed a tight upper bound, since some residual back-off may be needed, as shown by the CCDF curves presented in figure 42. Given that the percentage reduction observed in the required back-off is shown to be mainly a function of the number of stages, a more accurate gain estimation is given by

$$G_{MM} = \kappa \times PAPR_{RRC} - \overline{P}_{MM\_loss} \ , \tag{5.67}$$

where typical values for κ (i.e. the overall back-off reduction gain[7]) are given in table 7.

Number of Stages	1	2	3
% Reduction (κ)	80% ~ 95%	90% ~ 100%	95% 100%

Table 7: Typical observed reduction in required back-off when using MPMM.

5.2.3 *Bounding the Average Transmitted Power*

There are two simple methods of decreasing the peak power of bandwidth limited pulse shaped signal $x[n]$ in order to avoid HPA saturation. One, which consists of applying some back-off to the transmitted signal, i.e. by *scaling* it, has already been mentioned. This preserves spectrum shape as desired but badly degrades the system power efficiency. The opposite is to deliberately *clip* the signal before amplification [63, 67]. This generates serious spectral spreading but minimizes the drop in the signal's average power and therefore the PAPR of the transmitted signal, thus maximizing the system's power efficiency. The average power loss due to clipping is thus a lower limit on the average power penalty caused by the use of any peak power control technique. The comparison of power loss due to MM with the clipping method gives us some idea of how well the new technique performs, since it does not cause spectral spreading.

Let $x_c[n]$ be the signal resulting from the clipping operation. We will assume polar clipping, since MM performs polar distortion of the signal. The clipping signal is thus given by

$$|x_c[n]| = \begin{cases} A & , |x[n]| > A \\ |x[n]| & , |x[n]| \leqslant A \end{cases} \ , \tag{5.68}$$

$$\measuredangle x_c[n] = \measuredangle x[n] \ . \tag{5.69}$$

7 These typical values were obtained from the results presented in section 4.3.

By simple observation of figure 43, and remembering that we have been assuming $G_{RRC} = 1$, hence $\sigma_x^2 = \sigma_{s_{MM}}^2$, it easy to conclude that

$$\sigma_{s_{MM}}^2 \leqslant \sigma_{s_{x_c}}^2 \Leftrightarrow E\left\{|s_{MM}[n]|^2\right\} \leqslant E\left\{|x_c[n]|^2\right\} . \tag{5.70}$$

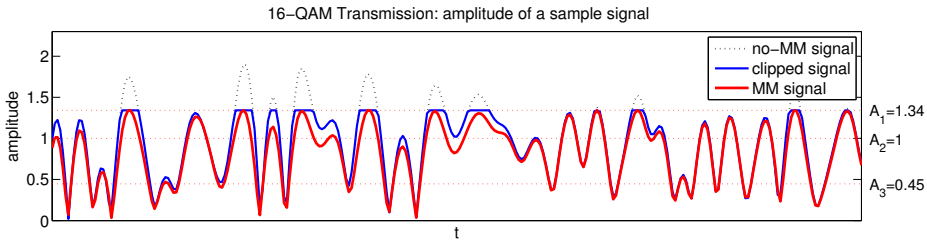


Figure 43: Sample of a transmitted 16-QAM RRC filtered signal with roll-off= 0.1 when the MPMM and clipping techniques for peak power control are used.

In order to compute the average power of the clipped signal, let us represent $x[n]$ in terms of its amplitude and phase components, i.e.

$$x[n] = x^I[n] + j\, x^Q[n] = R_x[n]\, e^{j\theta_x[n]} , \tag{5.71}$$

and let $p_{R_x}(r)$ denote the pdf of its amplitude distribution. The average power of the clipped signal is thus given by [63, 67],

$$P_{x_c} = E\left\{|x[n]|^2\right\} = \int_0^A r^2 p_{R_x}(r)\, dr + A^2 \int_A^{+\infty} p_{R_x}(r)\, dr . \tag{5.72}$$

Usually, assuming the validity of the central limit theorem, $x^I[n]$ and $x^Q[n]$ are considered to be independent Gaussian random processes, with distribution $\mathcal{N}\left(0, \sigma_x^2/2\right)$, and so the magnitude,

$$R_x[n] = \sqrt{(x^I[n])^2 + (x^Q[n])^2} , \tag{5.73}$$

follows a Rayleigh distribution [84] with

$$p_{R_x}(r) = \frac{2r}{\sigma_x^2} e^{-\frac{r^2}{\sigma_x^2}} \qquad \text{for } r \geqslant 0 , \tag{5.74}$$

and so (5.72) is expressed as,

$$P_{x_c} = \sigma_x^2 \left(1 - e^{-\frac{A^2}{\sigma_x^2}}\right) . \tag{5.75}$$

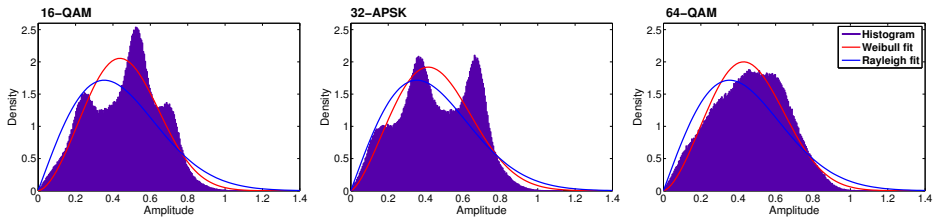


Figure 44: Envelope amplitude histogram of a non-MM bandwidth limited SC signal for different modulation formats, considering $\varepsilon_s = 1$ and RRC with $G_{RRC} = 1$ and roll-off 0.2. The Rayleigh and Weibull pdfs obtained through maximum likelihood estimation and that best fit the histogram are also drawn.

Although this model was found to be quite accurate for describing the amplitude of the transmitted signal of a multicarrier case, it does not work as well for a single carrier constellation unless the number of symbols M is very high [111]. For the SC case the pdf $p_{R_x}(r)$ was found to depend on the constellation geometry, as shown in figure 44, and on the RRC's filter roll-off. As an example, the results from statistical analysis curve fitting displayed in figure 44 show that $p_{R_x}(r)$ is more closely modeled by a Weibull distribution [84] than by Rayleigh, when considering 0.2 roll-off RRC filtering. The result (5.75) will be used even so as reference, since it is a well known result [63, 67, 109].

Numerical evaluation of (5.72) was thus performed for different M-ary constellations, using simulation data. The results are presented in figure 45 and compared with the Rayleigh approximation (5.75). The power of the MM signal given by (5.59) is also shown for a 3-stage MPMM system. Interestingly, the Rayleigh distribution is shown to provide a good approximation of the power of an MM signal. In addition, we may conclude that (5.75) constitutes a lower bound on P_{x_c} for the SC case.

The power efficiency gain, i.e. the reduction in the transmitted signal's PAPR, can also be observed in figure 45 as the difference with respect to the scaling method's curve (pure back-off). G_{MM} rises for higher order constellations because only outer constellation symbols suffer more severe distortion due to MM, which leads to a smaller fall in the average transmitted power. The same conclusion would be drawn from (5.61) and (5.66) considering MM factors statistical distributions.

Clipping implies a loss in average signal power of less than 1dB. The MPMM performs very well and differs by less than 1dB from the clipping case, showing that the MPMM can provide almost maximum power efficiency for a given constellation and RRC filter, while avoiding spectral regrowth. This is emphasized when comparing the transmitted average power using the MPMM technique with scaling, where the significant increase obtained in the system's power efficiency becomes clear, especially at sharper roll-offs.

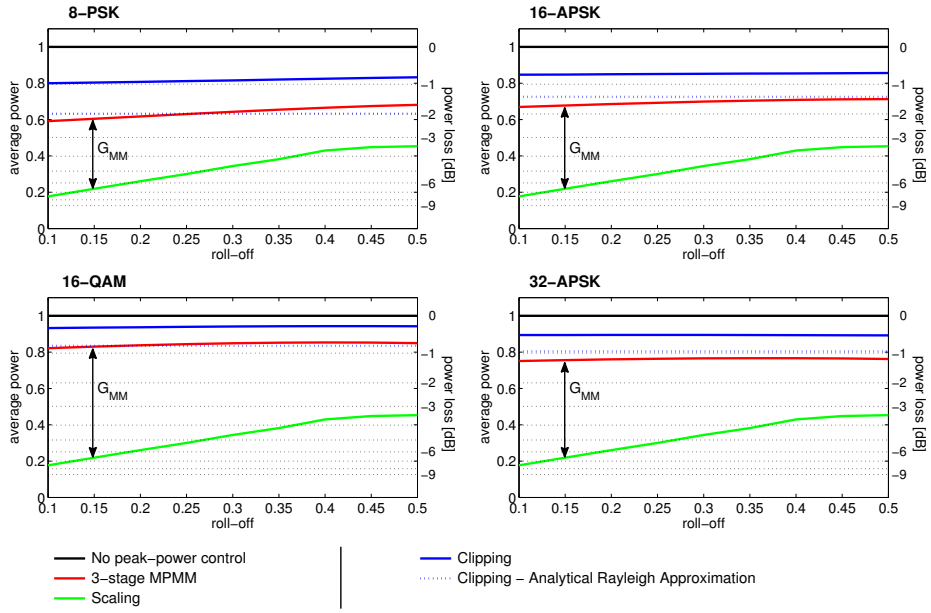


Figure 45: Transmitted signal's average power and correspondent power losses for different constellations and peak-power reduction techniques, considering $\varepsilon_s = 1$ and $G_{RRC} = 1$.

5.3 EVM AND SER ANALYSIS

Error vector magnitude (EVM) is a typical figure of merit used to evaluate accuracy of a modulation scheme when subjected to magnitude and phase distortions [5, 24, 109]. Requirements on EVM are already part of most wireless communication standards such as the 3GPP25.141 [1] and the IEEE802.16e-2005 Wimax standard [49].

Other popular figures of merit used to assess performance of digital modulation schemes are SNR, SER and BER. All these figures relate to each other, albeit in a non-trivial manner. Efforts have been made in order to get analytical close expressions among them in different contexts [5, 24, 86, 98].

This section presents an analytical study of the EVM and SER arising from MM distortion. Assuming a noiseless transmission, we derive expressions for the analytical computation of the EVM and bounds on the SER for the MM's blind reception and detection based on the MM EAC. We also address the case of an MM blind receiver that uses *automatic gain control* (AGC) [15, 43, 109] to adjust the amplitude of the constellation used in detection. The optimum scale factor to apply to the original constellation according to a criterion of minimizing the EVM is deduced.

5.3.1 *EVM Analytical Analysis*

Basically, EVM is a measure of the magnitude of the error vector between the measured received symbols and the expected ones. Since we are considering a noiseless transmission, the only source of distortion is due to MM and so the *error vector signal* $e_{MM}[n] \in \mathcal{C}$ can be computed as

$$e_{MM}[n] = s_{MM}[n] - \tilde{s}[n] \ , \tag{5.76}$$

where $\tilde{s}[n]$ is the reference signal (expected signal).

In order to make it independent of the constellation type, EVM is usually expressed as a percentage of the average power of the reference signal [24, 98, 109] and thus defined as

$$\mathrm{EVM}_{rms} = \sqrt{\frac{E\left\{|e_{MM}[n]|^2\right\}}{E\left\{|\tilde{s}[n]|^2\right\}}} = \sqrt{\frac{E\left\{|e_{MM}[n]|^2\right\}}{\frac{1}{M}\sum_{i=1}^{M}\left|\tilde{S}_i\right|^2}} \ , \tag{5.77}$$

where $\tilde{S}_i$ are the symbols of the reference M-ary constellation $\tilde{\mathcal{X}}$ used by the receiver on detection.

MM's blind detection

An MM blind receiver uses the original constellation $\mathcal{X}$ employed at the transmitter as reference. Hence, the error signal is given by[8]

$$e_{MM}[n] = s_{MM}[n] - s[n] = s[n]\,(m[n] - 1) \ . \tag{5.78}$$

Therefore, the power of the noise (i.e. error vector) due to MM is

$$\begin{aligned}\sigma^2_{e_{MM}} &= E\left\{|e_{MM}[n]|^2\right\} = \sum_{i=1}^{M}\left(|S_i|^2 \Pr(S_i)\int_0^1 (m-1)^2\, p(m \mid S_i)\, dm\right) \\ &= \sum_j \left(\frac{n_{A_j} A_j^2}{M}\int_0^1 (m-1)^2\, p\left(m \mid \mathbf{S}^{A_j}\right) dm\right) \ ,\end{aligned} \tag{5.79}$$

and so, considering (5.56) we get

$$\sigma^2_{e_{MM}} = \sum_j \left(\frac{n_{A_j} A_j^2}{M} \times \frac{\left(\beta_{A_j}+1\right)\beta_{A_j}}{\left(\alpha_{A_j}+\beta_{A_j}+1\right)\left(\alpha_{A_j}+\beta_{A_j}\right)}\right) . \tag{5.80}$$

8 Remember that we have assumed, without loss of generalization, that $G_{RRC} = 1$.

The average power of the original constellation is ε_s, and so EVM experienced under MM's blind detection can be evaluated analytically as

$$\mathrm{EVM}_{rms} = \sqrt{\sum_j \left(\frac{n_{A_j} A_j^2}{M\varepsilon_s} \times \frac{\left(\beta_{A_j}+1\right)\beta_{A_j}}{\left(\alpha_{A_j}+\beta_{A_j}+1\right)\left(\alpha_{A_j}+\beta_{A_j}\right)} \right)} . \tag{5.81}$$

Table 8 presents the results of the analytical computation (5.81) for α_{A_j} and β_{A_j} provided in table 5. The results refer to different normalized constellations, i.e. $\varepsilon_s = 1$, and multirate pulse shaping filters with different roll-offs and $G_{RRC} = 1$, when considering a 3-stage MPMM configuration at the transmitter.

We observe that the values computed from (5.81) follow the simulation results accurately, which again confirms the usefulness of the statistical model proposed in section 5.2.1. In addition, we also observe that the EVM experienced at the MM's blind detection is considerable, and is more severe for the APSK constellation. This agrees with the highest sensitivity to MM noise of the APSK constellations observed in sections 3.2.5 and 4.4.1, mainly because of their low $\mathrm{PAPR}_{\mathrm{Const}}$ (i.e. having constellation symbols closer to each other).

Detection using the MM EAC

As stated in section 4.4.2, the optimum detection of received symbols should take into account the distortion introduced by MM. Therefore, it was proposed to perform demodulation based on the EAC, where the average position of a constellation symbol $S_i \in \mathcal{X}$ after being MM is given by (4.50). Substituting (5.54) into (4.50) we get

$$\overline{S}_{i_{MM}} = S_i E\left\{ m \mid S_i \in \mathbf{S}^{A_j} \right\} = S_i \overline{m}_{A_j} , \tag{5.82}$$

	Analytical Model			Simulation		
Roll-off →	0.1	0.2	0.3	0.1	0.2	0.3
8-PSK	26.9%	24.0%	21.7%	26.9%	24.0%	21.7%
16-APSK	22.4%	20.3%	18.6%	22.4%	20.3%	18.5%
16-QAM	14.4%	12.8%	11.6%	14.5%	12.6%	11.4%
32-APSK	18.1%	16.4%	15.3%	18.0%	16.3%	15.1%
64-QAM	9.6%	8.1%	7.1%	9.2%	7.6%	6.6%

Table 8: EVM experienced at detection by an MM blind receiver. Assumptions are $\varepsilon_s = 1$, $G_{RRC} = 1$ and a 3-stage MPMM configuration at the transmitter.

whereby (5.56), $\overline{m}_{A_j}$ is

$$\overline{m}_{A_j} = E\Big\{ m \mid S_i \in \mathbf{S}^{A_j} \Big\} = \frac{\alpha_{A_j}}{\alpha_{A_j} + \beta_{A_j}} \,. \tag{5.83}$$

When the detection process uses the EAC the error vector signal is computed as

$$e_{MM}[n] = \big(m[n] - E\{m[n] \mid s[n]\} \big) s[n] \;, \tag{5.84}$$

and so its power is given by

$$\begin{aligned} \sigma^2_{e_{MM}} &= \sum_{i=1}^{M} \left(|S_i|^2 \Pr(S_i) \int_0^1 \big(m - E\{m \mid S_i\} \big)^2 p(m \mid S_i) \, dm \right) \\ &= \sum_j \left(\frac{n_{A_j} A_j^2}{M} \int_0^1 \Big(m - \overline{m}_{A_j} \Big)^2 p\Big(m \mid \mathbf{S}^{A_j} \Big) \, dm \right) , \end{aligned} \tag{5.85}$$

whereby (5.56) and (5.83), it follows that

$$\sigma^2_{e_{MM}} = \sum_j \left(\frac{n_{A_j} A_j^2}{M} \times \frac{\alpha_{A_j} \beta_{A_j}}{\Big(\alpha_{A_j} + \beta_{A_j} + 1 \Big) \Big(\alpha_{A_j} + \beta_{A_j} \Big)^2} \right) . \tag{5.86}$$

The average power of the EAC is determined as

$$\sigma^2_{EAC} = \frac{1}{M} \sum_{i=1}^{M} \left| \overline{S}_{i_{MM}} \right|^2 = \sum_j \left(\frac{n_{A_j} A_j^2}{M} \times \frac{\alpha_{A_j}^2}{\Big(\alpha_{A_j} + \beta_{A_j} \Big)^2} \right) , \tag{5.87}$$

which means that the EVM for detection using the EAC can be analytically evaluated as

$$EVM_{rms} = \sqrt{ \frac{ \sum_j \left(n_{A_j} A_j^2 \alpha_{A_j} \beta_{A_j} \Big/ \left[\Big(\alpha_{A_j} + \beta_{A_j} + 1 \Big) \Big(\alpha_{A_j} + \beta_{A_j} \Big)^2 \right] \right) }{ \sum_j \left(n_{A_j} A_j^2 \alpha_{A_j}^2 \Big/ \Big(\alpha_{A_j} + \beta_{A_j} \Big)^2 \right) } } \,. \tag{5.88}$$

The numerical evaluation of (5.88) is presented in table 9 and compared with the simulation results. We find that (5.88) provides a precise estimation of EVM

in accordance with the simulation, which once again confirms the correctness of the proposed statistical model.

But the most relevant conclusion that can be drawn from the results is that the use of the EAC leads to a significant decrease of the EVM experienced at reception which agrees with the conspicuous improvement in BER performance reported in section 4.4.2.

	Analytical Model			Simulation		
Roll-off →	0.1	0.2	0.3	0.1	0.2	0.3
8-PSK	15.2%	11.9%	9.6%	15.7%	12.3%	9.6%
16-APSK	14.2%	11.5%	9.4%	14.4%	11.6%	9.2%
16-QAM	10.5%	8.5%	7.1%	10.5%	8.2%	6.4%
32-APSK	12.1%	9.9%	8.3%	12.2%	9.8%	7.9%
64-QAM	8.2%	6.5%	5.3%	7.6%	5.7%	4.4%

Table 9: EVM experienced at detection when using MM EAC. Assumptions are $\varepsilon_s = 1$, $G_{RRC} = 1$ and a 3-stage MPMM configuration at the transmitter.

MM's blind detection using AGC

Optimum detection usually implies the use of an AGC at the receiver for synchronization purposes [9, 15] and to improve demodulation [43, 87], by amplitude scaling of the reference constellation used in detection.

Let $\eta \in \mathbb{R}^+$ denote the scaling factor to apply to the original constellation by an MM blind receiver using AGC. The error vector signal is thus

$$e_{MM}[n] = s_{MM}[n] - \eta\, s[n] = s[n]\,(m[n] - \eta)\ , \tag{5.89}$$

The scaling factor η must be chosen so that EVM is minimized. Substituting (5.89) in EVM definition (5.77) we get

$$\begin{aligned} \text{EVM}_{rms} &= \sqrt{\frac{E\left\{(m[n]-\eta)^2\,|s[n]|^2\right\}}{E\left\{\eta^2\,|s[n]|^2\right\}}} \\ &= \sqrt{\frac{\varepsilon_s\eta^2 - 2E\left\{m[n]\,|s[n]|^2\right\}\eta + \sigma^2_{s_{MM}}}{\varepsilon_s\eta^2}}\ , \end{aligned} \tag{5.90}$$

where we used (5.52) and (5.57).

Equation (5.90) is minimized when

$$\eta = \frac{\sigma^2_{s_{MM}}}{E\left\{m[n]\,|s[n]|^2\right\}} \tag{5.91}$$

whereby, combining equations (5.90) and (5.91), EVM becomes

$$\mathrm{EVM}_{rms} = \sqrt{1 - \frac{\left[E\left\{m[n]\,|s[n]|^2\right\}\right]^2}{\varepsilon_s\,\sigma^2_{s_{MM}}}}\,. \tag{5.92}$$

The term $E\left\{m[n]\,|s[n]|^2\right\}$ can be computed using (5.56), where, by manipulation we arrived at

$$E\left\{m[n]\,|s[n]|^2\right\} = \sum_j \left(\frac{n_{A_j} A_j^2}{M} \times \frac{\alpha_{A_j}}{\alpha_{A_j} + \beta_{A_j}}\right). \tag{5.93}$$

Table 10 presents the results both from the evaluation of (5.92) and from simulation. They agree with one another. The use of AGC according to a criterion of minimizing the EVM significantly improves the performance of an MM blind receiver, as can be concluded from comparing table 10 and table 8. Since, in practice, some sort of AGC is always used at synchronization and detection [9, 15, 43, 87], the EVM experienced by an MM blind receiver will always be much lower than reported in table 8.

However, even when employing AGC, the results are still worse than when using the MM EAC for detection (table 9), which proves the effectiveness of the EAC approach.

5.3.2 *SER Analytical Analysis*

Assume that a symbol $S_i \in \mathcal{X}$ is sent and that it is MM by m. As before, let $\tilde{\mathcal{X}}$ denote the reference constellation used by the receiver on detection, and $\tilde{S}_i$ the symbol matching S_i. Consider that transmission is noiseless. When using ML minimum distance detection[9] at the receiver, a symbol error occurs if

$$\exists \tilde{S}_j \in \tilde{\mathcal{X}} \setminus \{\tilde{S}_i\} : \; \left|mS_i - \tilde{S}_j\right| \leqslant \left|mS_i - \tilde{S}_i\right| \,. \tag{5.94}$$

9 See sections 2.3.1 and 2.3.3.

	Analytical Model			Simulation		
Roll-off →	0.1	0.2	0.3	0.1	0.2	0.3
8-PSK	15.0%	11.8%	9.6%	15.5%	12.2%	9.5%
16-APSK	14.3%	11.7%	9.6%	14.5%	11.8%	9.4%
16-QAM	11.5%	10.0%	9.0%	11.6%	9.8%	8.5%
32-APSK	13.0%	10.9%	9.4%	13.0%	10.8%	9.1%
64-QAM	8.8%	7.4%	6.5%	8.3%	6.8%	5.8%

Table 10: EVM experienced at detection by an MM blind receiver that uses AGC. As in table 8, the assumptions are $\varepsilon_s = 1$, $G_{RRC} = 1$ and a 3-stage MPMM configuration at the transmitter.

Solving (5.94) in order to m provides an upper bound m^{S_i} on the MM factor which, when applied to S_i, causes a detection symbol error. This value is given by

$$m^{S_i} = \min_{\tilde{S}_j \in \tilde{\mathcal{X}} \setminus \{\tilde{S}_i\}} \frac{1}{2} \times \frac{\left|\tilde{S}_i\right|^2 - \left|\tilde{S}_j\right|^2}{\left\langle S_i \cdot \tilde{S}_i \right\rangle - \left\langle S_i \cdot \tilde{S}_j \right\rangle}, \tag{5.95}$$

where $\langle S_a \cdot S_b \rangle$ is the scalar vector product[10].

Hence, SER experienced at reception can be computed explicitly as

$$\mathrm{SER} = \sum_{i=1}^{M} \left(\Pr(S_i) \int_0^{m^{S_i}} p(m \mid S_i)\, dm \right). \tag{5.96}$$

MM's blind detection

For MM's blind detection $\tilde{\mathcal{X}} = \mathcal{X}$, and so (5.95) simplifies to

$$m^{S_i} = \min_{S_j \in \mathcal{X} \setminus \{S_i\}} \frac{1}{2} \times \frac{|S_i|^2 - |S_j|^2}{|S_i|^2 - \langle S_i \cdot S_j \rangle}, \tag{5.97}$$

The m^{S_i} obtained for all $S_i \in \mathbf{S}^{A_j}$ will not all be equal. They depend on the constellation geometry. Table 11 presents, for MM's blind detection, the upper bound $m_{UB}^{A_j}$ and the lower bound $m_{LB}^{A_j}$ in the MM factors that cause symbol

10 A complex number $a + jb$ can be interpreted as a vector with coordinates (a, b), and so the scalar vector product as applied to complex numbers is defined as $\langle (a + jb) \cdot (c + jd) \rangle \triangleq ac + bd$.

errors when applied to any $S_i \in \mathbf{S}^{A_j}$. A symbol error occurs whenever $m < m_{LB}^{A_j}$ and there will be no error if $m > m_{UB}^{A_j}$.

We can thus define two analytical bounds on SER that are particularly useful when analyzing large constellations. Assuming that all symbols are transmitted with equal probability, and paying heed to (5.54) and (5.56), SER is bounded as

$$\sum_j \left(\frac{n_{A_j}}{M} \int_0^{m_{LB}^{A_j}} p\left(m \mid \mathbf{S}^{A_j}\right) dm \right) \leqslant \text{SER} \leqslant \sum_j \left(\frac{n_{A_j}}{M} \int_0^{m_{UB}^{A_j}} p\left(m \mid \mathbf{S}^{A_j}\right) dm \right). \quad (5.98)$$

Table 12 presents the SER simulation results for an MM blind receiver, as well as the evaluation of the analytical theoretical bounds. The 8-PSK case is omitted because the polar MM does not cause symbol errors when it is applied to constant amplitude constellations even though it increases the constellation's sensitivity to noise (EVM is non-null).

Again, the adopted statistical model is found to be accurate. The defined analytical bounds are quite tight and establish a narrow interval for the SER experienced through simulation.

Analysis of table 12 reinforces the conclusions drawn previously about the performance of MM's blind detection. In fact a relatively high SER is observed. In addition, we confirm that constellations with low $\text{PAPR}_{\text{Const}}$ are more prone to MM noise.

		$m_{LB}^{A_j}$	$m_{UB}^{A_j}$			$m_{LB}^{A_j}$	$m_{UB}^{A_j}$
8-APSK	A_1	0	0		A_1	0	0
16-APSK	A_1	0	0		A_2	0.6667	0.6667
	A_2	0.6201	0.6587		A_3	0.6667	0.6667
16-QAM	A_1	0	0		A_4	0.8	0.8
	A_2	0.6667	0.6667	64-QAM	A_5	0.8	0.8
	A_3	0.6667	0.6667		A_6	0.8	0.8571
32-APSK	A_1	0	0		A_7	0.8571	0.8571
	A_2	0.6302	0.6761		A_8	0.8571	0.8571
	A_3	0.7400	0.7694		A_9	0.8571	0.8571

Table 11: Limits on MM factors that cause symbol errors considering MM's blind detection. Given $S_i \in \mathbf{S}^{A_j}$, MM causes a symbol error whenever $m < m_{LB}^{A_j}$ and no error occurs if $m > m_{UB}^{A_j}$.

	Analytical lower bound			Simulation			Analytical upper bound		
Roll-off →	0.1	0.2	0.3	0.1	0.2	0.3	0.1	0.2	0.3
16-APSK	5.8	2.8	1.0	6.9	3.9	1.5	8.9	5.1	2.3
16-QAM	2.6	1.4	0.68	2.8	1.3	0.43	2.6	1.4	0.68
32-APSK	9.1	6.6	4.5	10.4	7.8	5.2	12.0	9.5	7.2
64-QAM	5.3	4.0	3.2	5.9	4.1	3.1	5.8	4.3	3.4

$(\cdots) \times 10^{-2}$

Table 12: SER experienced at detection by an MM blind receiver, when considering $\varepsilon_s = 1$, $G_{RRC} = 1$ and a 3-stage MPMM configuration at the transmitter.

EAC and MM's blind detection with AGC

A similar derivation can be followed for the cases of detection using the MM EAC or MM's blind detection using AGC. Considering reference constellations with symbols given by (5.82) for the EAC case, or the original constellation scaled by (5.91) for the AGC case, we can get analogous tables on m^{A_j} from the combination with (5.95). Then we can obtain, through analytical computation, an interval for the SER experienced in each instance. But the correctness of our analytical model has already been proved and no new insight would result therefrom. Hence, only the simulation results on SER are presented in table 13.

	Detection using EAC			MM's blind det. with AGC		
Roll-off →	0.1	0.2	0.3	0.1	0.2	0.3
16-APSK	2.3	0.92	0.13	2.2	0.86	0.093
16-QAM	1.4	0.50	0.081	1.5	0.56	0.14
32-APSK	4.1	2.4	1.1	4.0	2.3	0.97
64-QAM	3.8	2.4	1.3	4.2	2.8	1.9

$(\cdots) \times 10^{-2}$

Table 13: Simulation results for SER experienced at detection for: (a) receiver using the MM EAC on detection; (b) MM's blind detection using AGC. As in table 12, assumptions are $\varepsilon_s = 1$, $G_{RRC} = 1$ and a 3-stage MPMM configuration at the transmitter.

We observe a significant decrease in SER for both approaches compared with the MM's blind detection. Detection based on the MM EAC performs similarly to the AGC case for PSK constellations, but considerably better for QAM constellations. These results agree with the much lower EVM experienced when using the MM EAC, especially for QAM modulations. In addition, also note that SER is much lower for QAM than for PSK transmission, even when comparing constellations of considerably different sizes, e.g. 64-QAM and 32-APSK. This explains the higher BER performance losses observed in section 4.4 for PSK constellations.

5.4 PERFORMANCE ANALYSIS USING CHANNEL CODING

In this section we derive bounds on the BLER performance of the MM scheme over an AWGN channel when using FEC based on linear block coding. Hard- and soft-decision cases are both handled.

The development closely follows that used in section 2.3. Let $\mathcal{C}$ be a binary (n, k) linear block code with minimum Hamming distance d_{min} and weight enumerating polynomial $A(z)$. We also assume coherent BPSK transmission and that MM EAC is used in detection.

As in section 2.3, let $\mathbf{c} = [c_0, c_1, \cdots, c_{n-1}]$ denote a codeword $\in \mathcal{C}$, and $\mathbf{s} = [s_0, s_1, \cdots, s_{n-1}]$ be the corresponding BPSK sequence obtained from (2.41). Finally, let $\mathbf{s}_{MM} = \left[s_{MM_0}, s_{MM_1}, \cdots, s_{MM_{n-1}}\right]$ denote the sequence that results from applying MM to $\mathbf{s}$, and $\mathbf{r} = [r_0, r_1, \cdots, r_{n-1}]$ the n sampled outputs of the receiver's matched filter. When MM is used, each received sampled output may be expressed as[11]

$$r_i = s_{MM_i} + n_i = m_i \sqrt{E_s}\,(2c_i - 1) + n_i\,, \tag{5.99}$$

where m_i and n_i are r.v. representing the MM factor applied to symbol s_i and the AWGN noise added to the MM transmitted symbol s_{MM_i}, with $n_i \sim \mathcal{N}(0, N_0/2)$.

5.4.1 *Hard-Decision Decoding Performance*

The beta distributions of the MM factors that apply to each of the BPSK constellation symbols ($\sqrt{E_s}$ and $-\sqrt{E_s}$) are known to be the same according to (5.56). Hence, assuming that transmitted symbols are equiprobable, by reasoning similar to that followed in section 2.3.2, it is easy to conclude that when performing hard-decision decoding according to (2.50), the optimum decision threshold is

11 Throughout this chapter the average energy per transmitted symbol has been denoted as ε_s. But in section 2.3.2 the energy of a transmitted BPSK symbol was denoted as E_s. Since the development in this section closely follows that in section 2.3.2, the former notation will be used, and so, in the BPSK context, whenever there is a reference to a previous current chapter's formula $\varepsilon_s = E_s$.

also at $\tau = 0$. The system from the modulator's input to the hard-decoder's output can thus be seen as a BSC channel with a crossover probability given by (2.53), i.e.

$$p = \Pr\left(r_i < 0 \,\middle|\, s_i = \sqrt{E_s}\right) = \Pr\left(r_i > 0 \,\middle|\, s_i = -\sqrt{E_s}\right) \ . \tag{5.100}$$

Although MM does not cause any symbol errors in the BPSK sent symbols, because $m_i \in]0, 1]$, the constellation's increased sensitivity to noise should be taken into account when computing the crossover probability p of the equivalent BSC channel. Hence, equation (5.100) can be rewritten as

$$p = \int_0^1 \Pr\left(n_i > m\sqrt{E_s} \;\middle|\; s_{MM_i} = -m\sqrt{E_s}\right) p\left(m \mid s_i = \sqrt{E_s}\right) dm \ , \tag{5.101}$$

and so, from combining (5.54) and the Gaussian distribution of n_i with (5.101), we get

$$p = \int_0^1 Q\left(m\sqrt{\frac{2E_s}{N_0}}\right) \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\,\Gamma(\beta)} m^{\alpha-1} (1-m)^{\beta-1} \, dm \ . \tag{5.102}$$

The integral may be expressed as

$$p = \frac{1}{2} + \overline{m}\sqrt{\frac{E_s}{\pi N_0}} \times {}_3F_3\left(\begin{matrix} \frac{1}{2} \ , & 1+\frac{\alpha}{2} \ , & \frac{1}{2}+\frac{\alpha}{2} \\ \frac{3}{2} \ , & 1+\frac{\alpha+\beta}{2} \ , & \frac{1}{2}+\frac{\alpha+\beta}{2} \end{matrix} \ ; -\frac{E_s}{N_0} \right) \ , \tag{5.103}$$

where $\overline{m}$ is the expectation of the MM factors given by (5.83), and ${}_pF_q$, with $p, q \in \mathbb{N}^+$, is the *generalized hypergeometric* function [2, 23], defined as

$${}_pF_q\left(\begin{matrix} a_1, a_2, \cdots, a_p \\ b_1, b_2, \cdots, b_q \end{matrix} \ ; z \right) = \sum_{n=0}^{\infty} \frac{(a_1)_n (a_2)_n \cdots (a_p)_n}{(b_1)_n (b_2)_n \cdots (b_q)_n} \frac{z^n}{n!} \ , \tag{5.104}$$

with $(a)_n$ denoting the *Pochhamer symbol*, which can be computed as

$$(a)_n = a\,(a+1)\cdots(a+n-1) = \frac{\Gamma(a+n)}{\Gamma(a)} \ . \tag{5.105}$$

As in section 2.3.2, it is convenient to express p as a function of the signal to noise ratio per information bit E_b/N_0. Relation (2.56) has to be modified according to the fact that MM is now being used. Since it is assumed that all symbols occur with equal probability, the average energy of a symbol of the MM EAC, i.e. $E_{s_{MM}}$, equals the average transmitted power of the MM signal, which

is given by (5.60) for constant amplitude constellations such as BPSK. Hence, E_b relates to E_s according to

$$E_b = \frac{nE_{s_{MM}}}{k} = \frac{E_{s_{MM}}}{r_c} = \frac{E\{m^2\}\,E_s}{r_c}\,, \tag{5.106}$$

which makes it possible, through simple substitution in (5.103), to express p as a function of E_b/N_0.

It has been shown in section 2.3.2 that for a transmission over a BSC with crossover probability p, the decoder's probability of committing a decoding error is upper bounded by (2.64). By combining (2.64) with p given by (5.103) we thus obtain an upper bound on P_E when performing hard decoding for a system using MM. Again, equality holds in equation (2.64) if the linear block code is perfect, e.g. the Golay $(23, 12)$ code.

Analytical and simulation results under hard-decision decoding for this Golay code are plotted in figure 46. The results refer to a system using a 3-stage MPMM configuration at the transmitter that performs RRC pulse shaping with roll-off$= 0.2$. Hard-decision simulation results when no MM is used, presented in figure 7, previously, are also reproduced for comparison.

This analytically-derived bound is found to be exact, just as expected. When it is compared with the non-MM case the graph shows a fall of only 1dB in performance in the probability of error domain at $P_E = 10^{-5}$. This again proves the high efficiency of the MM scheme, since, for the tested configuration, the need to back-off is completely eliminated, which is equivalent to a gain of about 5.8dB (see table 2 and figure 31).

5.4.2 *Soft-Decision Decoding Performance*

The computation of the exact error probability P_E for soft-decision decoding is not a trivial task, as was shown in section 2.3.3. We have therefore resorted to the union bound (2.74). The same approach will be taken in the study on the MM case. As in section 2.3.3, an upper bound on P_E can be computed as

$$P_E \leqslant \sum_{\bar{c}_i} \left(\Pr[\bar{c}_i] \sum_{c_j \in \mathcal{C} \setminus \{\bar{c}_i\}} \Pr\big[C(\mathbf{r}, \mathbf{c}_j) > C(\mathbf{r}, \bar{c}_i)\big] \right)\,, \tag{5.107}$$

where $\bar{c} \in \mathcal{C}$ denotes the codeword assumed to be sent.

The MM EAC is assumed to be used on detection. Hence, equation (5.99) can be rewritten as the sum of the expected MM BPSK symbol and the noise contributions from the MM distortion and the AWGN channel, i.e.

$$\begin{aligned} r_i &= \overline{m} s_i + (m_i - \overline{m})\, s_i + n_i \\ &= \overline{m}\sqrt{E_s}\,(2c_i - 1) + e_{MM_i}^{(EAC)} + n_i \,, \end{aligned} \tag{5.108}$$

where $e_{MM_i}^{(EAC)}$ is the error vector signal (5.84). The two noise sources[12], e_{MM_i} and n_i, are independent and both have zero mean. The variance of e_{MM_i} is given by (5.86), which for BPSK simplifies to

$$\sigma_{e_{MM}}^2 = \sigma_m^2 E_s = \frac{\alpha\beta}{(\alpha+\beta+1)(\alpha+\beta)^2} E_s \,, \tag{5.109}$$

where σ_m^2 denotes the variance of the MM factors. As stated before, it is considered that $n_i \sim \mathcal{N}(0, N_0/2)$.

Remember that although the MM system has memory because of the treatability reasons of the problem it has been assumed that m_i only depends on the current constellation symbol s_i being MM, and so the MM noise contributions, e_{MM_i} with $i = 0, \cdots, n-1$, that add to each symbol s_i of a codeword $\mathbf{s} \in \mathcal{C}$ can be considered independent of one another. Furthermore they all obey the same statistical distribution. The composite channel (MM+AWGN) can thus be seen as symmetric, and since the code is linear the sum $\sum_{\mathbf{c}_j \in \mathcal{C}\setminus\{\bar{\mathbf{c}}_i\}} \Pr\big[C(\mathbf{r}, \mathbf{c}_j) > C(\mathbf{r}, \bar{\mathbf{c}}_i)\big]$ is the same for all $\bar{\mathbf{c}}_i$. Therefore, just like in section 2.3.3, it can be assumed that the all-zero codeword $\bar{\mathbf{c}}_0$ is transmitted, and computation of (5.107) also reduces to (2.74), i.e.

$$P_E \leqslant \sum_{\mathbf{c}_j \in \mathcal{C}\setminus\{\bar{\mathbf{c}}_0\}} \Pr\big[C(\mathbf{r}, \mathbf{c}_j) > C(\mathbf{r}, \bar{\mathbf{c}}_0)\big] \,. \tag{5.110}$$

Expanding the probability that $C(\mathbf{r}, \mathbf{c}_j) > C(\mathbf{r}, \bar{\mathbf{c}}_0)$ by using equation (5.108), we get the same equation that was obtained in (2.75), that is

$$\begin{aligned} \Pr\big[C(\mathbf{r}, \mathbf{c}_j) > C(\mathbf{r}, \bar{\mathbf{c}}_0)\big] &= \Pr\left[\sum_{i=0}^{n-1} r_i \left(c_{ji} - \bar{c}_{0i}\right) > 0\right] \\ &= \Pr\big[S_j > 0\big] \,, \end{aligned} \tag{5.111}$$

where S_j denotes the sum $\sum_{i=0}^{n-1} r_i \left(c_{ji} - \bar{c}_{0i}\right)$.

The random variables r_i, with $i = 0, \cdots, n-1$, are no longer Gaussian. But they are still independent as a result of noise contributions e_{MM_i} and n_i, being independent not only of each other (i.e. for the same index i), but also from

12 Hereafter, we will omit superscript on e_{MM_i}. We only used it in (5.108) to make it clear that we were referring to the error vector signal for the detection case using the MM EAC.

symbol to symbol. It thus follows from (5.108) that the mean and variance of r_i are

$$m_{r_i} = E\{r_i\} = -\overline{m}\sqrt{E_s} \,, \tag{5.112}$$

$$\sigma^2_{r_i} = E\Big\{\big(r_i - m_{r_i}\big)^2\Big\} = \sigma^2_m E_s + \frac{N_0}{2} \,. \tag{5.113}$$

Hence, S_j is a sum of independent random variables formed by $w_H(\mathbf{c}_j)$ non-null terms that correspond to the positions where $\mathbf{c}_j$ differs from $\bar{\mathbf{c}}_0$. By the central limit theorem [84], S_i has a cumulative distribution function that approaches a normal (Gaussian) distribution with mean

$$m_{S_j} = -w_H(\mathbf{c}_j)\ \overline{m}\sqrt{E_s} \,, \tag{5.114}$$

and variance

$$\sigma^2_{S_j} = \left(\sigma^2_m E_s + \frac{N_0}{2}\right) w_H(\mathbf{c}_j) \ , \tag{5.115}$$

and so probability (5.111) is given by

$$\Pr\big[S_j > 0\big] = Q\left(-\frac{m_{S_j}}{\sigma_{S_j}}\right) = Q\left(\overline{m}\sqrt{\frac{w_H(\mathbf{c}_j)\, E_s}{\sigma^2_m E_s + \frac{N_0}{2}}}\right) \ . \tag{5.116}$$

Combining (5.116) with (5.106), we can express $\Pr\big[S_j > 0\big]$ as a function of E_b/N_0. After proper manipulation we obtain

$$\Pr\big[S_j > 0\big] = Q\left(\overline{m}\sqrt{\frac{w_H(\mathbf{c}_j)}{\sigma^2_m + \frac{E\{m^2\}}{2r_c\frac{E_b}{N_0}}}}\right) \ . \tag{5.117}$$

The union bound on the probability of decoding error P_E can thus be obtained straightforwardly through substitution of (5.117) in (5.110), which gives

$$P_E \leqslant \sum_{w=d_{min}}^{n} A_w Q\left(\overline{m}\sqrt{\frac{w_H(\mathbf{c}_j)}{\sigma^2_m + \frac{E\{m^2\}}{2r_c\frac{E_b}{N_0}}}}\right) \ . \tag{5.118}$$

where A_w is the w-th order coefficient of the weight enumerating polynomial $A(z)$ of the linear code $\mathcal{C}$.

As in section 2.3.3, a very good approximation of (5.118) that does not require the knowledge of $A(z)$ can be obtained for large values of E_b/N_0. Here, the term relating to codewords with minimum Hamming weight d_{min} is dominant in the summation (5.118), and so P_E can be upper bounded by

$$P_E \leqslant A_{d_{min}} Q\left(\overline{m}\sqrt{\frac{d_{min}}{\sigma_m^2 + \frac{E\{m^2\}}{2r_c\frac{E_b}{N_0}}}}\right) . \tag{5.119}$$

Bounds given by equations (5.118) and (5.119) are compared in figure 46, with the exact word error probability (obtained by Monte Carlo simulation), for the case of the binary (23, 12) Golay code. The results refer to the same system configuration used in the previous section for the hard-decision case.

We observe that both bounds are indeed quite tight, providing a very good approximation of the exact word error probability under soft-decision decoding. This not only validates the usefulness of the derived expressions and the assumptions made during development, but it also confirms the correctness of the MM factors' statistical model.

Figure 46 also plots the soft-decision decoding error probability when no MM is used. As for the hard-decision case, performance loss at $P_E = 10^{-5}$ is less than 1dB, which proves the high efficiency of the system compared with the reduction of about 5.8dB obtained in the required back-off for the system configuration tested. In addition, figure 46 also shows that the $\approx$ 2dB difference between hard- and soft-decision remains unaltered.

Error Floor

For very high E_b/N_0 values, the MM noise contribution e_{MM_i} becomes dominant over the AWGN noise. In this case, probability (5.117) can be approximated by

$$Pr[S_j > 0] \approx Q\left(\overline{m}\sqrt{\frac{w_H(\mathbf{c}_j)}{\sigma_m^2}}\right) . \tag{5.120}$$

which does not depend on the signal noise ratio per information bit. Hence, an error floor under soft-decision decoding will be observed for high values of

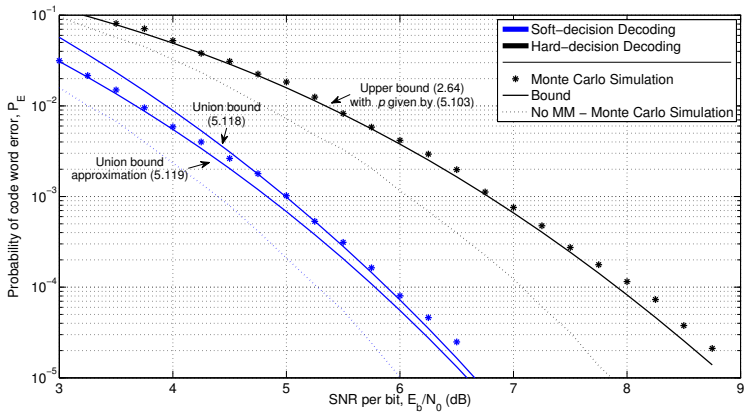


Figure 46: Comparison of bounds with exact error probability for hard-decision and soft-decision decoding for the Golay $(23, 12)$ code, when MM is used. Results refer to a system using a 3-stage MPMM configuration at the transmitter and an RRC pulse shaping filter with roll-off=0.2.

E_b/N_0. This is dictated by the MM distortion and it can be estimated based on the union bound (5.110), when using approximation (5.120), which gives

$$P_E\,(\textit{error floor}) \approx \sum_{w=d_{min}}^{n} A_w Q\left(\overline{m}\sqrt{\frac{w_H(\mathbf{c}_j)}{\sigma_m^2}}\right). \tag{5.121}$$

This error floor is fortunately extremely low. As an example, for the (23, 12) Golay code when considering the system configuration in figure 46 the computation of (5.121) yields 4.3485×10^{-56}.

5.5 CONCLUSION

A mathematical model based on the parameterized beta distributions has been proposed to characterize the statistical distributions of the MM factors for different SC constellations. Building on the proposed model, an analytical performance analysis of the MM technique was carried out in terms of: the power of the transmitted signal, the EVM and SER experienced in a noiseless transmission, and in the BLER when channel coding is used. The derived expressions and bounds have been shown to be quite accurate. This not only allows us to confirm the exactness and usefulness of the proposed statistical model, it also provides an analytical validation as to the effectiveness of the MM technique as a highly efficient peak-power reduction technique that does not cause spectral spreading.

6

CONCLUSIONS AND FUTURE WORK

6.1 CONCLUSIONS

This thesis has addressed the problem of reducing the PAPR of band limited SC signals. We were seeking to reduce the power amplifier back-off on communication systems such as VSAT, whose capacity is limited by the low power efficiency of the transmitter's non-linear HPA.

Research was focused on the development of the magnitude modulation technique, an effective PAPR reduction method first proposed (in the nineties) by Miller *et al.* [74, 75], and later enhanced by Tomlinson *et al.* [4, 108] with a proposal of a low-complexity LUT approach. However, the effectiveness of their approaches was limited to constant amplitude constellations, i.e. PSK modulations.

This work has made significant contributions to the MM concept. The LUT approach has been improved considerably and a new, highly efficient, polyphase MM algorithm has been proposed. Both techniques can now be applied to high-order non-constant amplitude constellations. However, the affordable treatment of the huge number of symbol combinations to be handled when applying the LUT's method still limits, in this case, the constellation's size to $M \leqslant 16$.

Other important achievements include the development of detection techniques that substantially reduce the noise sensitivity due to MM, and the proposal of a mathematical model for the statistical distributions of the MM factors that permits the analytical performance evaluation of MM systems.

In more detail, the main contributions of this thesis are:

- The LUT-MM approach has been improved and extended to constellations up to $M \leqslant 16$ in size [32, 34, 38]. In order to cope with the additional complexity, new clipping limits have been proposed [32] and the method's ability to avoid phase modulation improved [34]. By exploring the constellation's symmetry and the RRC's linear phase characteristic we have also been able to present a new method for compacting MM look-up tables and speed-up their computation. We have further contributed with a detailed analysis of the different schemes to compute the MM tables.

- A new, highly efficient, *Multistage Polyphase Magnitude Modulation* algorithm for PAPR reduction, has been developed, where MM factors are

computed in real time by a low-complexity polyphase filter system [35–38]. The main advantages of this technique have been highlighted. It has been proved that MPMM is independent of the modulation mapping geometry, making it possible to extend the MM concept to higher order and/or non-constant amplitude constellations. In fact, the analysis of the MPMM system has demonstrated that there are no constraints on the type of its input signal and, as noted, any number of MPMM base-system blocks can be added in a serial cascade, with each additional stage providing a fine control of the envelope excursion of the band-limited signal to be transmitted. The MPMM's efficiency is demonstrated by showing that it provides a better control of the envelope's peak power of SC band limited signals than previous MM approaches, even when only considering 1-stage. Furthermore, with just a few stages (typically 3), it is possible to almost eliminate the need for back-off, even at very low filter roll-off values.

- To offset the increased sensitivity to noise caused by MM, two effective techniques for improving the detection of the MM signal have been presented, viz.: detection based on the MM estimated average constellation (EAC) [36]; and detection with estimation of MM factors in the receiver. These techniques were developed with the aim of ameliorating the correct estimation of the a priori code bit LLRs used in the soft-decision decoding of the MM signal. It has been shown that EAC detection can provide a notable improvement in performance with respect to MM's blind reception without entailing changes in traditional receivers. The main innovation in the second technique is the combination of MM factors' estimation and soft decoding in a feedback loop scheme. It was shown that this method is particularly efficient for higher order constellations with $M \geqslant 32$ (more sensitive to MM distortion), especially when less powerful FEC coding is used.

- A mathematical model based on parameterized beta distributions has been proposed for characterizing the statistical distributions of MM factors for different SC constellations [38, 39]. This work also contains an analytical performance analysis of the MM technique in terms of the power of the transmitted signal, the EVM and SER experienced in a noiseless transmission, and the BLER when channel coding is used. Based on the proposed model, closed analytical expressions were provided for the evaluation of the EVM, as well as tight bounds on the SER and BLER experienced. Model correctness was confirmed by comparison with experimental results, and the derived expressions and bounds were found to be quite accurate.

The effectiveness of the proposed techniques was validated by means of simulation. It considered constellations up to $M = 64$ in size and pulse shaping with sharp roll-offs ($\alpha \leqslant 0.35$). Performance was evaluated in terms of the

reduction observed in the PAPR, the required back-off to drive the HPA close to saturation, and in the BER experienced over the AWGN channel when using a FEC system based on LDPC codes. As a measure of overall performance, it was defined the net back-off gain (3.34).

It should be stressed that a significant net back-off gain above 5dB (at BER$= 10^{-5}$) was obtained with the MPMM scheme using a $(1248, 832)$ LDPC code [49]). This confirms the extremely high efficiency of the new MPMM technique proposed in this thesis. Although, net_G_{BO} has also shown some dependency on the error correcting capabilities of the selected FEC scheme, the methods developed for improving the detection of the MM signal proved that it is possible to obtain effective high net_G_{BO} gains for different combinations of modulation and coding.

Finally, the importance of the developed analytical model for the statistical characterization of MM factors' distributions should be emphasized. It provides an analytical validation of the effectiveness of the MM technique as a highly efficient peak-power reduction technique.

6.2 FUTURE WORK

Power efficiency is a key issue in our world today. The interest in PAPR reduction in telecommunications proves it. The problem extends over different (but related) fields such as SC and MC communications and *software radio* (SWR). This work has addressed a small piece of this problem, although it provides a solid starting basis for future research.

One open subject is the study of MM applied to SC-FDMA [95]. The recent interest in such an SC modulation scheme, thanks to its PAPR characteristics being better than those of OFDM, makes it a promising research topic. However, the PAPR problem is still present in SC-FDMA, and it is our belief that MM could be a good solution to minimize it.

Adapting the MM concept to the MC case and comparing it with state-of-the art techniques is another demanding challenge.

Another field in which the PAPR problem arises, and which is also of interest to us and the subject of future investigation, is SWR. An SWR signal combines multiple standards where each has its own system parameters. Some standards use MC modulation (like OFDM based systems) whereas others use SC modulation (like GSM). Since the SWR signal is a combination of these different SC and MC modulation based standards it may well inherit high PAPR.

APPENDIX

A

NYQUIST FILTER AND ITS PAPR CONTRIBUTION

As stated in chapter 1, the largest contribution to the PAPR of a band limited SC signal comes from the Nyquist pulse shaping. Thus it makes sense to briefly discuss the Nyquist filter characteristics and the type of filters normally used to implement it, such as the root-raised cosine filter.

In this section we also evaluate the contribution of the Nyquist filter to the total PAPR of a bandwidth limited SC modulated signal, i.e. the PAPR of signal $x[n]$ for the generic SC transmitter system in figure 1 (chapter 1).

A.1 NYQUIST PULSE SHAPING

The Nyquist pulse is the signal pulse required to communicate over band-limited channels with no ISI. The continuous-time version of this family of filters is described in almost every communications book [10, 65, 87]. However, only a few discuss their digital version counterpart that is generally used in practice [44, 105]. For a pulse duration of L samples (interpolator up-sampling factor [44, 76] in figure 1), the digital pulse-shaping filter should satisfy the so-called *Nyquist condition for zero ISI*, i.e. the impulse response h_{Nyq} is such that

$$h_{Nyq}[kL] = 0 \quad , \forall k \in \mathbb{Z}\backslash\{0\} \ . \tag{A.1}$$

A frequently used class of filters verifying the Nyquist criteria is the *raised cosine* RC filters. The frequency response of an RC filter in the continuous-time domain is

$$H_{RC}(j\omega) = \begin{cases} T_{SYMB} & , \ 0 \leqslant |\omega| \leqslant \frac{(1-\alpha)\pi}{T_{SYMB}} \\ \frac{T_{SYMB}}{2}\left\{1 + \cos\left[\frac{T_{SYMB}}{2\alpha}\left(|\omega| - \frac{(1-\alpha)\pi}{T_{SYMB}}\right)\right]\right\} & , \ \frac{(1-\alpha)\pi}{T_{SYMB}} \leqslant |\omega| \leqslant \frac{(1+\alpha)\pi}{T_{SYMB}} \\ 0 & , \ |\omega| \geqslant \frac{(1-\alpha)\pi}{T_{SYMB}} \end{cases} \tag{A.2}$$

where T_{SYMB} is the symbol's duration time period and α is the filter roll-off which specifies the excess bandwidth with respect to $1/(2T_{SYMB})$ (the minimum

bandwidth required, according to Nyquist theory [10, 87], to transmit a sequence at rate $1/T_{SYMB}$).

In order to maximize the *signal-to-noise ratio* (SNR) at the output of the receiver's front-end the receiver filter must be matched to the transmitting shaping filter [87]. Hence, the overall RC spectral characteristic is usually split evenly between the filters used at the transmitter and receiver, i.e. both having a spectra of magnitude $|H_{RRC}(j\omega)| = \sqrt{|H_{RC}(j\omega)|}$, which is why this filter is called *root-raised cosine*.

The impulse response of a digital RRC is obtained by sampling the continuous time impulse response[1] $h_{RRC}(t) = \mathcal{F}^{-1}\{H_{RRC}(j\omega)\}$, at a rate $f_{SMPL} = L/T_{SYMB}$, which gives,

$$h_{RRC}[n] = \frac{4\alpha n^2 \cos\left[\frac{\pi n(1+\alpha)}{L}\right] + L\sin\left[\frac{\pi n(1+\alpha)}{L}\right]}{\left[1-\left(4\alpha\frac{n}{L}\right)^2\right]\pi n\sqrt{L}}\ , \tag{A.3}$$

which was scaled, to ensure $\sum_n |h[n]|^2 = 1$. The autocorrelation of the RRC impulse response, i.e. $h_{RRC}[n] * h^*_{RRC}[-n]$, is a Nyquist filter and thus satisfies (A.1).

The *oversampling factor* $L \in \mathbb{N}$ determines the up-sampling factor that precedes the pulse shaping filter in figure 1. According to the Nyquist sampling theorem [76, 82], in order to avoid aliasing we must guarantee that

$$f_{SMPL} \geqslant \frac{2}{T_{SYMB}}(1+\alpha)\ . \tag{A.4}$$

In practice, the use of a higher up-sampling factor imposes fewer restrictions on the design of the DAC's reconstruction filter, allowing a considerable reduction in its order and so lowering its cost [44]. Therefore, it is usual to operate the digital RRC filter at 4 or 8 samples per symbol, i.e. to set $L = 4$ or 8.

To limit the impulse response $h_{RRC}[n]$ to a finite duration, $h_{RRC}[n]$ is usually truncated by applying a rectangular window that spreads over $2N+1$ symbols[2]. The parameter N should be large enough to guarantee almost zero ISI [44]. To solve this problem and also the fact that the side-lobes in the spectrum of the finite duration $h_{RRC}[n]$ are quite high (rectangular windowing), other windows are also used [44] or, alternatively, filter coefficients are optimized [11, 105, 113].

1 $F^{-1}\{H(j\omega)\}$ stands for the inverse Fourier transform of $H(j\omega)$ [82].
2 This leads to a filter of $2NL+1$ taps

A.2 NYQUIST FILTER CONTRIBUTION TO SIGNAL PAPR

Consider the generic block diagram of a typical SC transmitter, as shown in figure 1 (section 1.2). It is our intention to prove that, as stated in chapter 1, the PAPR of a single carrier band limited signal $x[n]$ is

$$\mathrm{PAPR}_{x[n]} = \mathrm{PAPR}_{\mathrm{Const}} + \mathrm{PAPR}_{\mathrm{RRC}} \ , \tag{A.5}$$

with $\mathrm{PAPR}_{\mathrm{Const}}$ and $\mathrm{PAPR}_{\mathrm{RRC}}$ given by (1.3) and (1.4), respectively.

Symbols at the output of the demodulator are assumed to be independent and identically distributed (i.i.d.) random variables (r.v.). As shown before, the transmitted average power is given by

$$\sigma_s^2 = E\left\{|s[n]|^2\right\} = \frac{1}{M}\sum_{i=1}^{M}|S_i|^2 = \varepsilon_s \ , \tag{A.6}$$

and so the expression (1.3) for $\mathrm{PAPR}_{\mathrm{Const}}$, i.e. the PAPR of $s[n]$, is readily yielded from applying the PAPR definition when using (A.6).

Let us now prove that the contribution from Nyquist pulse shaping is given by (1.4). Consider the more general case of a zero mean WSCS signal $s[n]$ with average power σ_x^2. Its PAPR is expressed as

$$\mathrm{PAPR}_{s[n]} = 10\log_{10}\left(\frac{\max|s[n]|^2}{\sigma_s^2}\right) \quad (\mathrm{dB}) \ . \tag{A.7}$$

The proof makes use of the theoretical insights addressed in section 5.2. From theorem 5.1, since $s[n]$ is WSCS, and the autocorrelation of $h[n]$ satisfies the Nyquist condition for zero ISI, then the average power of the band limited signal $x[n]$ is given by

$$\sigma_x^2 = \frac{\sigma_s^2}{L}\sum_{n=-\infty}^{\infty}|h[n]|^2 \ . \tag{A.8}$$

In order to obtain the PAPR of signal $x[n]$, we just have to evaluate

$$\max_n|x[n]| = \max_n\left|\sum_k h[k]\,s_u[n-k]\right| \ . \tag{A.9}$$

Since the transmitter system is multirate the pulse-shaping can be performed at symbol rate according to polyphase decomposition (4.15), as presented in

figure 25 (see section 4.1.3). Denoting the output signal of each polyphase filter $E_i(z)$ as $y_i[n]$, we get

$$\begin{aligned}\max_n |x[n]| &= \max_{i=0,1,\cdots,L-1} |y_i[n]| \\ &= \max_{i=0,1,\cdots,L-1} \left| \sum_k h[kL+i]\, s_u[n-k] \right| , \end{aligned} \tag{A.10}$$

which, according to algebra properties, can be upper bounded as

$$\begin{aligned}\max_n |x[n]| &\leqslant \max_{i=0,1,\cdots,L-1} \sum_k \left| h[kL+i]\, s_u[n-k] \right| \\ &\leqslant \max_n |s[n]| \times \max_{i=0,1,\cdots,L-1} \sum_k \left| h[kL+i] \right| . \end{aligned} \tag{A.11}$$

The PAPR of $x[n]$ can thus be computed as

$$\begin{aligned}\mathrm{PAPR}_{x[n]} &= 10\log_{10}\left(\frac{\max|x[n]|^2}{\sigma_x^2}\right) \\ &= 10\log_{10}\left(\frac{\max_n |s[n]| \times \max_{i=0,1,\cdots,L-1} \sum_k |h[kL+i]|}{\frac{\sigma_s^2}{L}\sum_{n=-\infty}^{\infty} |h[n]|^2}\right) \end{aligned} \tag{A.12}$$

which can be written as

$$\mathrm{PAPR}_{x[n]} = \underbrace{10\log_{10}\left(\frac{\max_n |s[n]|}{\sigma_s^2}\right)}_{\mathrm{PAPR}_{s[n]}} + \underbrace{10\log_{10}\left(\frac{\max_{i=0,1,\cdots,L-1} \sum_k |h[kL+i]|}{\frac{1}{L}\sum_{n=-\infty}^{\infty} |h[n]|^2}\right)}_{\mathrm{PAPR}_{\mathrm{RRC}}} . \tag{A.13}$$

(q.e.d.)

B

WIMAX AND DVB-S2 LDPC CODES

Low-density parity-check codes LDPC [20, 21, 68] are currently one of the most important classes of linear block codes that are capable of approaching the Shannon limit [99]. Their importance is reflected by their recent inclusion in several standards, such as WiMAX IEEE 802.16e [49], *Mobile Broadband Wireless Access* (MBWA) IEEE 802.20 [50], *Ethernet* IEEE 802.3an [51] and the satellite's digital video broadcast DVB-S2 [19].

This was the main reason why we chose them for the FEC system of the testbed we used to evaluate the performance of peak power control techniques proposed in this work. The codes we used were taken from the WiMAX IEEE 802.16e standard [49] because of their linear encoding complexity and their affordable length for computer simulations. Their structure is therefore set forth in this appendix, though briefly.

This appendix also reports some research work on parallel architectures for encoding and decoding [30, 31, 33] the LDPC codes used by DVB-S2 that has been carried out alongside with the thesis' main work[1].

B.1 STRUCTURED LDPC-IRA CODES

LDPC codes [20, 21, 68] are linear block codes defined by sparse parity-check matrices $\mathbf{H}$ and they are usually represented by Tanner graphs [106]. A Tanner graph is a bi-partite graph formed by two types of nodes. Check nodes (v^C), one per each code constraint, and bit nodes, one per each codeword bit (information and parity, respectively, v^I and v^P), with the connection edges between them being given by $\mathbf{H}$.

They are decoded using low complexity iterative belief propagation algorithms operating over the Tanner graph description [12]. However, a major drawback is their high encoding complexity caused by the fact that the generator matrix $\mathbf{G}$ is, in general, not sparse. In order to overcome this problem, standards such as DVB-S2 and WiMAX have adopted a special class of LDPC codes, with linear encoding complexity, named *Irregular Repeat-Accumulate* (IRA) [58].

1 Their importance is shown by the high number of known citations to [31].

B.1.1 *IRA Codes*

Although the parity check matrix **H** of an LDPC code is sparse, the generator matrix **G** needed for encoding is usually not sparse, and this leads to storage and encoding complexity problems. An IRA code is characterized by a parity check matrix **H** of the form,

$$\mathbf{H}_{(n-k)\times n} = \left[\mathbf{A}_{(n-k)\times k} \mid \mathbf{B}_{(n-k)\times(n-k)}\right] = \left[\begin{array}{cccc|ccccc} a_{00} & a_{01} & \cdots & a_{0,k-1} & 1 & & & & \\ a_{10} & a_{11} & \cdots & a_{1,k-1} & 1 & 1 & & & \\ \vdots & \vdots & & \vdots & & \ddots & \ddots & & \\ a_{n-k-2,0} & a_{n-k-2,1} & \cdots & a_{n-k-2,k-1} & & & 1 & 1 & \\ a_{n-k-1,0} & a_{n-k-1,1} & \cdots & a_{n-k-1,k-1} & & & & 1 & 1 \end{array}\right]. \quad \text{(B.1)}$$

where **B** is a staircase lower triangular matrix. By restricting **A** to be sparse, an LDPC-IRA code is obtained [18].

The codes defined by (B.1) are systematic, i.e. the codewords have the following format, $\mathbf{c} = [\mathbf{i} \mid \mathbf{p}]$, with the information bits $\mathbf{i} = [i_0\, i_1 \cdots i_{k-1}]$ being associated with the **A** matrix, and the parity check bits $\mathbf{p} = [p_0\, p_1 \cdots p_{n-k-1}]$, with the **B** matrix. In terms of Tanner graph representation, this means that check nodes (CN) related to the code constraints are connected with two different types of bit nodes (BN): the *information nodes* (IN) that represent the information bits, and the *parity nodes* (PN) associated with the parity check bits, as shown in figure 47.

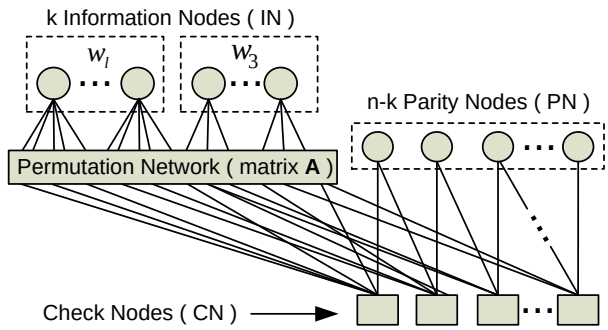


Figure 47: Tanner graph description of an LDPC-IRA code [60].

The bi-diagonal structure of the **B** sub-matrix corresponds to a zigzag connectivity between CNs and PNs, which allows the computation of the parity check bits directly from the **A** matrix, in a recursive manner, according to:

$$
\begin{aligned}
p_0 &= a_{00} i_0 \oplus a_{01} i_1 \oplus \cdots \oplus a_{0,k-1} i_{k-1} \\
p_1 &= a_{10} i_0 \oplus a_{11} i_1 \oplus \cdots \oplus a_{1,k-1} i_{k-1} \oplus p_0 \\
&\vdots \\
p_0 &= a_{n-k-1,0} i_0 \oplus a_{n-k-1,1} i_1 \oplus \cdots \oplus a_{n-k-1,k-1} i_{k-1} \oplus p_{n-k-2} \,,
\end{aligned} \tag{B.2}
$$

where $\oplus$ is the GF(2) addition operator. For LDPC-IRA codes, since the **A** matrix is sparse, this presents an O(n) encoding complexity [58].

B.1.2 *DVB-S2 LDPC Codes*

The **H** matrices of the DVB-S2 LDPC codes have properties other than being of IRA type [18, 19]. Some periodicity constraints were imposed on the pseudo-random design of the **A** matrices to make it possible a significant reduction in the storage requirement without code performance loss.

The matrix **A** construction technique is based on dividing the v^I nodes in disjoint groups of M consecutive ones. All the v^I nodes of a group l have the same weight w_l, and it is only necessary to specify the v^C nodes that connect to the first v^I of the group in order to know the v^C nodes that connect to each of the $M-1$ remaining v^I nodes.

Denote by $\{r_1, r_2, \cdots, r_{w_l}\}$ the indices of the v^C nodes that connect to the first v^I of group l. The indices of the v^C nodes that connect to v^I_i, with $0 \leqslant i \leqslant M-1$, of group l can be obtained by,

$$
\begin{aligned}
&\{ (r_1 + i \times q) \bmod (n-k), (r_2 + i \times q) \bmod (n-k), \cdots, \\
&\quad (r_{w_l} + i \times q) \bmod (n-k) \}
\end{aligned} \tag{B.3}
$$

with $q = (n-k)/M$ and $M = 360$ (a common factor for all DVB-S2 supported codes). The structure of **H** is presented in figure 48.

Another property of matrix **A** is that for each supported code there is a set of groups of v^I nodes of constant weight $w > 3$ (w is code dependent), while all the others have weight 3.

B.1.3 *WiMAX LDPC Codes*

The LDPC codes in WiMAX [8, 49] standards are quasi-cyclic codes. The standard specifies six different code classes spanning four different code rates. All defined codes are based on the expansion of a $\mathbf{H}'$ base matrix which has an IRA-type structure. The $\mathbf{H}'$ matrix consists of 24 columns and $(1-R)\times 24$ rows[2], with each integer entry $p \geqslant -1$, describing a z-by-z sub-matrix which is either a p-right shifted permuted identity matrix (for $p \geqslant 0$) or a zero matrix ($p = -1$).

2 R stands for the code rate.

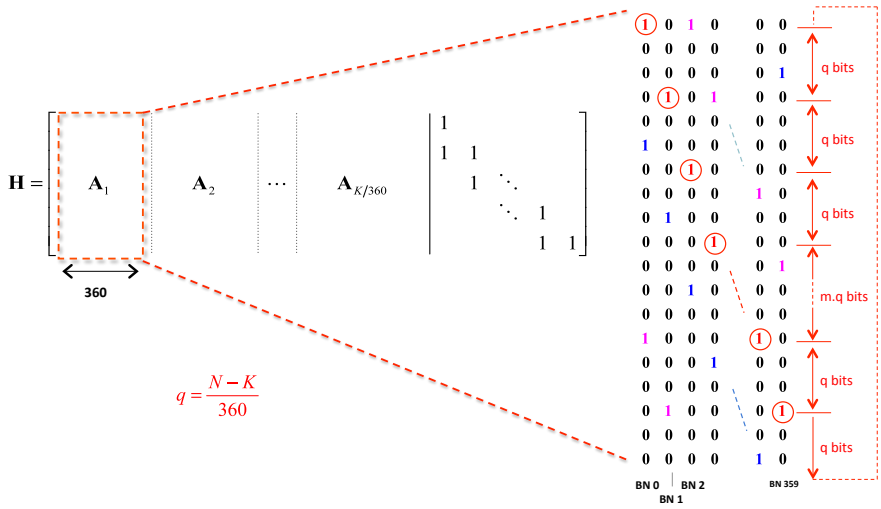


Figure 48: **H** matrix structure of a DVB-S2 LDPC-IRA code.

The first R×24 columns correspond to the message information and the rest $(1-R)\times 24$ to the parity information. The size of sub-matrices z-by-z is variable and ranges from 24×24 to 96×96 with a granularity of 4, therefore supporting 19 codeword sizes that ranges from $n = 576$ to $n = 2304$. The parity check matrix of the rate 1/2 code for $z = 96$ case is shown in figure 49 [8].

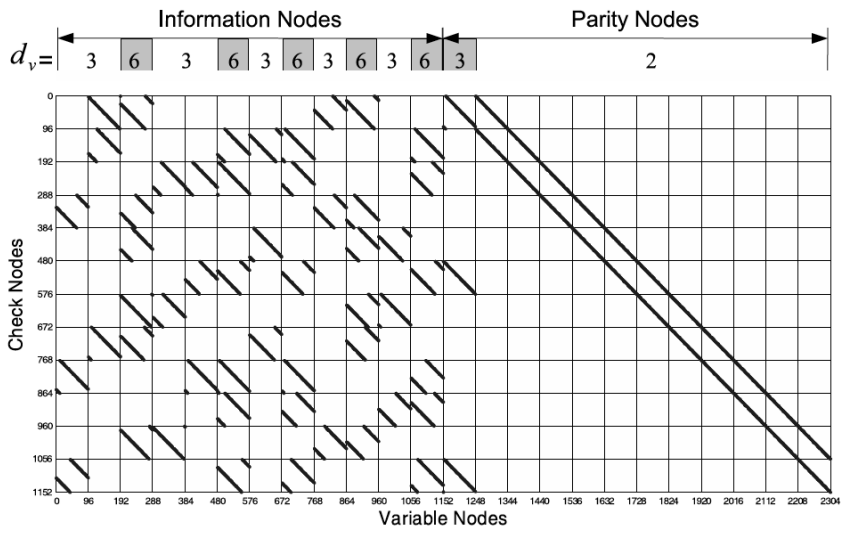


Figure 49: **H** matrix structure of a WiMAX LDPC code for rate 1/2 and $z = 96$ [8].

B.2 HARDWARE ARCHITECTURES FOR DVB-S2 LDPC-IRA CODES

Standards like DVB-S2 and WiMAX specify different combinations of modulation and coding (MODCOD) schemes. For instance, DVB-S2 supports two different frame lengths (16200 bits for low delay applications and 64800 bits otherwise) and a set of different code rates (1/4, 1/3, 2/5, 1/2, 3/5, 2/3, 3/4, 4/5, 5/6, 8/9 and 9/10) for both frame lengths and for different modulation schemes [18, 19]. A normally important issue is the design of the LDPC encoder and decoder architectures to be compliant with all MODCOD schemes.

This section presents efficient LDPC encoding and decoding hardware architectures [30, 31, 33] for DVB-S2 which were developed during this work.

B.2.1 *DVB-S2 LDPC Encoder Architecture*

By proper manipulation of (B.2) and considering the periodicity factor $M = 360$ common to all DVB-S2 LDPC IRA codes we developed an efficient encoder algorithm shared by all code options (length and rate) which is able to perform partial-parallel computation of the parity check bits, thereby increasing the encoder throughput.

Encoder algorithm

Let $IN(r)$ denote the indices of the information bits that participate in the parity check node restriction r, i.e. in v_r^C, and let $CN(c)$ denote the indices of CNs that are connected to v_c^I, where the letters "r" and "c" refer to row and column of $\mathbf{H}$. Hence, equations (B.2) can be rewritten as,

$$\begin{aligned} p_0 &= \bigoplus_{z \in IN(0)} i_z = S_0 \\ p_1 &= \bigoplus_{z \in IN(1)} i_z \oplus p_0 = S_1 \oplus S_0 \\ &\vdots \\ p_r &= S_r \oplus p_{r-1} = \bigoplus_{i=0}^{r} S_i \,, \end{aligned} \tag{B.4}$$

for $r = 0, \cdots, (n-k-1)$ and with,

$$S_r = \bigoplus_{z \in IN(r)} i_z \,. \tag{B.5}$$

The preceding equation can be recursively computed for each CN in the following way,

```
for r = 0 : (n − k − 1) do
    for each z ∈ IN(r) do
        S_r = S_r ⊕ i_z
    end
end
```
(B.6)

with S_r (for all r) being initialized to zero.

Instead of processing the CNs one by one, which requires the a priori knowledge of all information bits, we can swap the processing order from horizontal to vertical. So, for each new information bit received by the encoder the associated S_r values are updated according to,

```
for c = 0 : (k − 1) do
    for each r ∈ CN(r) do
        S_r = S_r ⊕ i_c
    end
end
```
(B.7)

Considering the DVB-S2 LDPC's codes periodicity, each inner cycle of (B.7) can be performed simultaneously for $M = 360$ CNs, thereby increasing the encoder throughput. According to (B.3), we know that for each group of M received information bits, say $\{i_c, i_{c+1}, \cdots, i_{c+M-1}\}$, with $c \bmod M = 0$, we only need to know $CN(c)$ in order to find $CN(c+j)$, for $j = 1, \cdots, M-1$. In fact, if v_r^C is connected to v_c^I (i.e. $r \in CN(c)$), then $[(r + j \times q) \bmod (n-k)] \in CN(c+j)$.

Hence, the iterative encoding procedure (B.7) can be rewritten as,

```
for c = 0 : M : (k − M) do
    for each r ∈ CN(r) do
        S_r                       = S_r                       ⊕ i_c
        S_{(r+q) mod (n−k)}       = S_{(r+q) mod (n−k)}       ⊕ i_{c+1}
          ⋮
        S_{(r+(M−1)×q) mod (n−k)} = S_{(r+(M−1)×q) mod (n−k)} ⊕ i_{c+M−1}
    end
end
```
(B.8)

If we store the S_r values as in the following binary matrix,

$$\mathbf{S} = \begin{bmatrix} S_0 & S_q & S_{2q} & \cdots & S_{359q} \\ S_1 & S_{q+1} & S_{2q+1} & \cdots & S_{359q+1} \\ S_2 & S_{q+2} & S_{2q+2} & \cdots & S_{359q+2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ S_{q-1} & S_{2q-1} & S_{3q-1} & \cdots & S_{n-k-1} \end{bmatrix}_{(q \times M)} \tag{B.9}$$

we can easily check that all updated S_r values in (B.8) are in the same row of $\mathbf{S}$. The only problem is that they may not be in the order needed to perform the iterative computation specified by (B.8). However, by observing the structure of $\mathbf{S}$ we conclude that we just need to perform a cycle shift of the corresponding $\mathbf{S}$ row or, alternatively, a rotation of the information vector. The second solution is preferable, since the operation on $\mathbf{S}$ can be easily vectorized in this way.

Considering the required row, $\mathbf{S}(r \bmod q, :)$, and the information vector, $\mathbf{i} = [i_c \; i_{c+1} \; \cdots \; i_{c+M-1}]$, the inner cycle (B.8) can thus be written as,

$$\mathbf{S}(r \bmod q, :) = \mathbf{S}(r \bmod q, :) \oplus \mathrm{rot}_{r \,\mathrm{div}\, q}(\mathbf{i}) \;, \tag{B.10}$$

where $\mathrm{rot}_a(\cdot)$ is a right circular shift of a bits.

Once the S_r values are known we can easily compute the parity bits, p_r, according to the recursive operations in (B.4). But this method would imply a significant encoder delay. Once again it is possible to increase the throughput by using vectorized approaches. By computing the checksum of all columns of $\mathbf{S}$ we obtain the following vector,

$$\mathbf{s} = \left[\bigoplus_{i=0}^{q-1} S_i \;\; \bigoplus_{i=q}^{2q-1} S_i \;\; \bigoplus_{i=2q}^{3q-1} S_i \;, \cdots, \bigoplus_{i=358q}^{359q-1} S_i \;\; \bigoplus_{i=359q}^{360q-1} S_i \right]_{(1 \times M)} \;, \tag{B.11}$$

which can be transformed by the following matrix operation in GF(2),

$$\mathbf{L} \times \mathbf{s}^T = \left[\bigoplus_{i=0}^{q-1} S_i \;\; \bigoplus_{i=0}^{2q-1} S_i \;\; \bigoplus_{i=0}^{3q-1} S_i \;, \cdots, \bigoplus_{i=0}^{359q-1} S_i \;\; \bigoplus_{i=0}^{360q-1} S_i \right] = \mathbf{s'} \;, \tag{B.12}$$

where $\mathbf{L}$ is an $M \times M$ lower triangular matrix of ones. Next, the vector $\mathbf{s'}$ undergoes a logical right shift of one bit in order to obtain,

$$\left[0 \;\; \bigoplus_{i=0}^{q-1} S_i \;\; \bigoplus_{i=0}^{2q-1} \;, \cdots, \bigoplus_{i=0}^{360q-1} S_i \right] = \left[0 \;\; p_{q-1} \;\; p_{2q-1} \;, \cdots, \;\; p_{359q-1} \right] \;. \tag{B.13}$$

Finally, for $j = 0, \cdots, q-1$, we calculate M parity bits at a time using the following procedure,

$$\left[p_j \; p_{q+j} \; p_{2q+j} \; , \cdots , \; p_{359q+j}\right] = \left[p_{j-1} \; p_{q+j-1} \; p_{2q+j-1} \; , \cdots , \; p_{359q+j-1}\right] \oplus \mathbf{S}(j,:) \; . \tag{B.14}$$

The sorting in natural order of the results in (B.14) can be included in the encoder interleaver design that follows the LDPC encoder [19].

Encoder architecture

Figure 50 presents the encoder architecture that implements the algorithm just described [30, 33]. It is composed of three major modules which are responsible for computing the S_r values (Task 1), performing the checksum of all columns of the **S** memory (Task 2) and, finally, computing the parity check bits (Task 3), according to (B.10), (B.11) and (B.14), respectively.

The proposed architecture is independent of the chosen DVB-S2 LDPC-IRA code. In fact the **S** RAM only has to be designed for the worst case scenario,

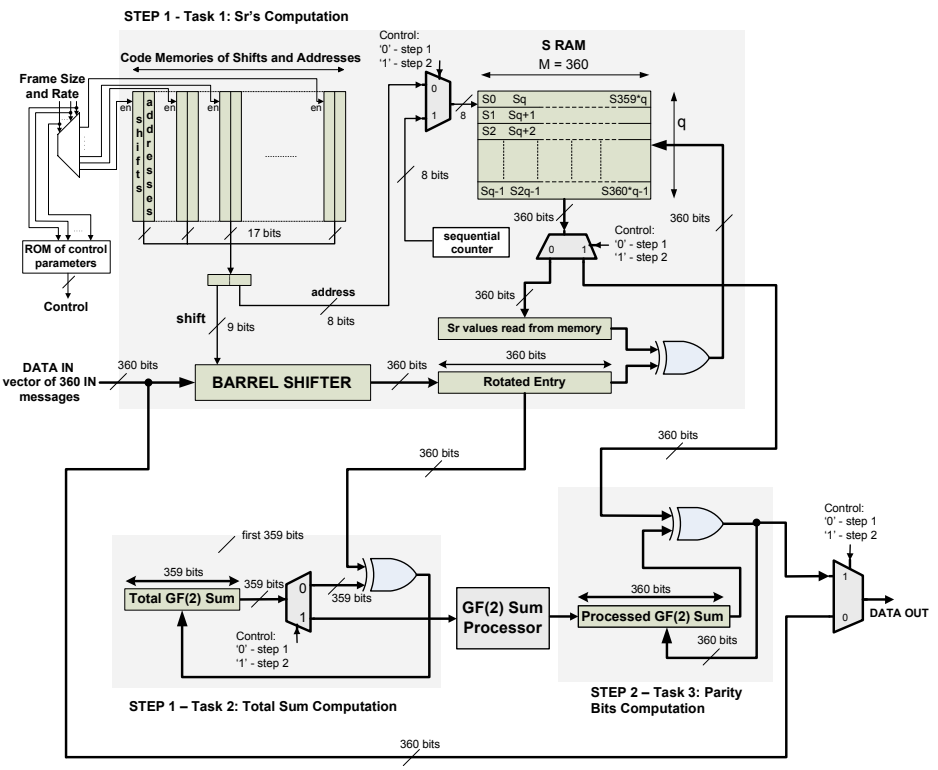


Figure 50: Architecture for the DVB-S2 LDPC-IRA encoder.

i.e. $q = 135$. Based on the code rate and frame length, the appropriate row of the control ROM is selected, where the encoder control parameters for each DVB-S2 LDPC code (the number of IN's groups of weight w) are stored. The corresponding ROM containing shift and address values is read sequentially in order to perform Task 1, according to (B.10) and (B.8). Address and shift values stored in ROMs can be easily computed, according to (B.10), from tables in the annexes B and C of DVB-S2 standard [19].

Task 2 and Task 1 are performed simultaneously. In fact, it is not necessary to wait until the end of S_r computation in order to obtain the checksum of all columns of the **S** memory. This would imply a longer delay. Instead, at each S_r computation the GF(2) addition of the current shifted information vector to the previous accumulated result is performed.

After Step 1, consisting of tasks 1 and 2, the first 359 values of the *Total GF(2) Sum* register are processed by a simple combinatorial module which produces (B.13). Once computed, this value is written into the *processed GF(2) sum* register and Task 3 is performed, computing 360 parity bits per clock cycle, according to (B.14).

B.2.2 *DVB-S2 LDPC Decoder Architecture*

The huge dimensions of the LDPC-IRA codes adopted by the DVB-S2 standard make the adoption of a fully parallel architecture that maps the Tanner graph structure impractical [28]. Furthermore, this solution is code dependent, which means that a different decoder is required for each code defined by the standard.

The best known solutions are based on highly vectorized partial parallel architectures [17, 60], that explore the particular characteristics of the DVB-S2 LDPC-IRA codes such as their periodic nature ($M = 360$) shared by them all. One solution was proposed in [60], whose architecture uses M functional units working in parallel. In this section we show that it is possible to reduce the number of functional units by any integer factor of M, without any addressing overhead and keeping its efficient memory mapping scheme unchanged. Our approach not only overcomes the disadvantages of the architecture described in [60], it is also more flexible and easily reconfigurable according to the decoder constraints.

M Parallel Processing Units Approach

The periodicity properties (B.3) of the DVB-S2 LDPC-IRA codes make it possible to simultaneously process the v^I and v^C node sets, whose indices are given by,

$$C^{(c)} = \{c,\ c+1, \cdots, c+M-1\} \quad \text{with,} \quad c \bmod M = 0\ , \tag{B.15}$$

and,

$$R^{(r)} = \{r,\ r+q,\ r+2q, \cdots, r+(M-1)\,q\} \quad \text{with,} \quad 0 \leqslant r \leqslant -1\ , \tag{B.16}$$

respectively, which significantly simplifies the decoder control. In fact, according to (B.3), assuming that $v^I_{\tilde{c}}$ is connected to v^C_r, then $v^C_{r+i\times q}$, with $0 \leqslant i \leqslant M-1$, is also connected to $v^I_{c+(\tilde{c}-c+i) \bmod M}$, where $c = M\times(\tilde{c} \operatorname{div} M)$ is the index of the first v^I of the group $C^{(c)}$ to which $v^I_{\tilde{c}}$ belongs.

The architecture shown in figure 51 is based on M functional units (FU) working in parallel with shared control signals [28], that process both v^C (in check mode) and v^I nodes (in bit mode) in a flooding schedule manner. As result of the zigzag connectivity between v^P and v^C nodes they are updated jointly in check mode following a horizontal schedule approach [100].

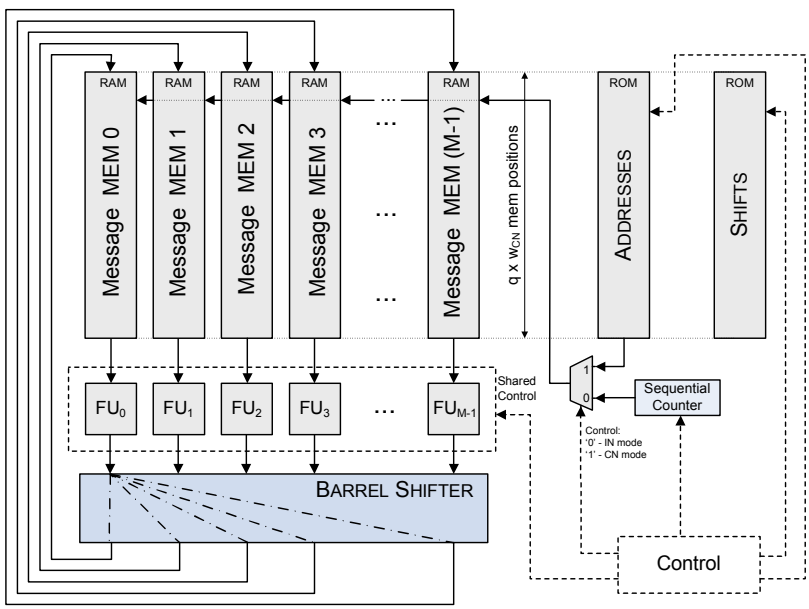


Figure 51: DVB-S2 LDPC decoding architecture using M parallel processing units [60].

Each FU unit is shared by a constant number of v^I, v^C and v^P nodes, depending on the code length and rate. More precisely, for a $(n,\ k)$ DVB-S2 LDPC IRA code the FU_i, with $0 \leqslant i \leqslant M-1$, updates sequentially in bit mode the $v^I_{\{i,i+M,i+2M,\cdots,i+(\alpha-1)M\}}$ nodes, with $\alpha = k/M$. In check mode the same FU updates the $v^C_{\{j,j+1,\cdots,j+q-1\}}$ and $v^P_{\{j,j+1,\cdots,j+q-1\}}$ nodes, with $j = i\times q$. This guarantees that when simultaneously processing the nodes in group $C^{(c)}$ the computed messages have as destination a set $R^{(r)}$, where each one is processed

by a different FU. Considering (B.3), the new computed messages only need to be right-rotated to be handled by the correct v^C nodes.

The same happens when processing each $R^{(r)}$ set where, according to (B.3), the right rotation must be reversed so that the new computed messages have as destination the exact v^I nodes.

The shuffling network (barrel shifter) is responsible for the correct message exchange between v^C and v^I nodes, emulating the code Tanner graph. The shift values stored in ROM (figure 51) can be easily obtained from annexes B and C of the DVB-S2 standard tables [19].

The messages sent along the Tanner graph edges are stored in RAM (see figure 51). If we adopt a sequential RAM access in bit mode then the access in check mode must be indexed, or vice-versa, with both options being valid. Without loss of generalization, let us assume sequential access in bit mode. Denoting by $\mathbf{r}_i = \left[r_{i_1} \; r_{i_2} \cdots \; r_{i_{w_i}} \right]$ the vector of v^C node indices connected to the v_i^I node, of weight w_i, the message memory mapping is specified by the following matrix,

$$\mathbf{R} = \begin{bmatrix} \mathbf{r_0} & \mathbf{r}_1 & \cdots & \mathbf{r}_{M-1} \\ \mathbf{r}_M & \mathbf{r}_{M+1} & \cdots & \mathbf{r}_{2M-1} \\ \vdots & \vdots & \vdots & \vdots \\ \mathbf{r}_{(\alpha-1)M} & \mathbf{r}_{(\alpha-1)M+1} & \cdots & \mathbf{r}_{\alpha M-1} \end{bmatrix}_{(q \times w_C) \times M} \tag{B.17}$$

where, w_C, is a code related parameter (v^C weight is $w_C + 2$, except for the first one (B.1)).

In order to process each $R^{(r)}$ set in check mode, the required memory addresses can be obtained by searching the matrix $\mathbf{R}$ rows where the index r appears.

Decoding Decomposition by a Factor of M

The simplicity of the shuffling mechanism and the efficient memory mapping scheme are the major strengths of the architecture just described [60]. But the large number of FUs and the long width of the barrel shifter require a huge silicon area. Since this architecture is able to provide a throughput far higher than the minimum mandatory rate of 90 Mbps we may reduce the number of FUs. In fact, we show that this can be done by any factor of M [31, 33].

Let $L, N \in \mathbb{N}$ be factors of M, with $M = L \times N$. A set $C^{(c)}$, as defined by (B.15), can be decomposed by subsampling in L subgroups as:

$$
\begin{aligned}
C_0^{(c)} &= \{c,\ c+L,\ c+2L,\ \cdots,\ c+(N-1)L\} \\
C_1^{(c)} &= \{c+1,\ c+1+L,\ c+1+2L,\ \cdots,\ c+1+(N-1)L\} \\
&\vdots \\
C_{L-1}^{(c)} &= \{c+L-1,\ c+2L-1,\ c+3L-1,\ \cdots,\ c+NL-1\}\ .
\end{aligned} \tag{B.18}
$$

According to (B.3), each sub-group $C_\gamma^{(c)}$, with $0 \leqslant \gamma \leqslant L-1$, can be described in terms of the first node of subgroup $v_{c+\gamma}^{I}$. If v_r^C is connected to the first information node of subgroup $C_\gamma^{(c)}$, then, $v^C_{(r+i\times L\times q)\bmod(n-k)}$ is connected to the i-th v^I node of the referred subgroup, considering that $0 \leqslant i \leqslant N-1$.

Similarly, $R^{(r)}$ can be decomposed in L sub-groups as:

$$
\begin{aligned}
R_0^{(r)} &= \{r,\ r+Lq,\ r+2Lq,\ \cdots,\ r+(N-1)Lq\} \\
R_1^{(r)} &= \{r+q,\ r+(L+1)q,\ \cdots,\ r+((N-1)L+1)q\} \\
&\vdots \\
R_{L-1}^{(r)} &= \{r+(L-1)q,\ r+(2L-1)q,\ \cdots,\ r+(NL-1)q\}\ ,
\end{aligned} \tag{B.19}
$$

and each subgroup $R_\beta^{(r)}$, with $0 \leqslant \beta \leqslant L-1$, can be described just in terms of the first element, $v^C_{r+\beta\times q}$. If $v_{\tilde{c}}^I$ is connected to the first node of sub-set $R_\beta^{(r)}$, then $v^I_{c+((\tilde{c}-c+i\times L)\bmod M)}$ with $c = M\times(\tilde{c}\ \mathrm{div}\ M)$ is connected to the i-th v^C of the considered subgroup, for $0 \leqslant i \leqslant N-1$.

Framework (B.18) and (B.19) shows that the subsampling approach preserves the key modulo M properties, and thus we conclude that subgroups $C_\gamma^{(c)}$ and $R_\beta^{(r)}$ can be processed individually using the same architecture as described in [60], but now considering only N processing units. In fact, when simultaneously processing a group $C_\gamma^{(c)}$, the computed messages have as destination a set $R_\beta^{(r)}$ and vice-versa. The new proposed architecture is shown in figure 52.

The sub-sampling strategy therefore allows a linear reduction (by a factor of L) of the hardware resources occupied by the FUs blocks, while it significantly reduces the complexity of the barrel shifter ($O(N\log_2 N)$) and simplifies the routing problem. Yet, at first glance, it may seem that this strategy implies an increase by L in the size of the system ROM (Shifts and Addresses). Fortunately, if we know the properties of subgroups $C_0^{(c)}$ and $R_0^{(r)}$ then we automatically know the properties of the remaining subgroups $C_\gamma^{(c)}$ and $R_\beta^{(r)}$, respectively (with $0 \leqslant \gamma,\ \beta \leqslant N-1$).

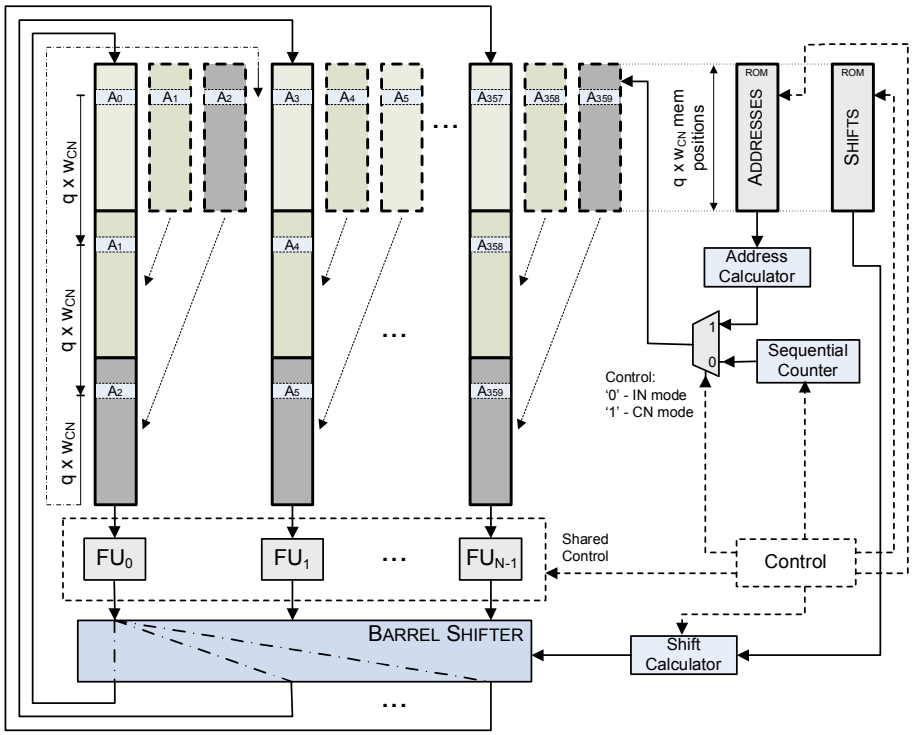


Figure 52: DVB-S2 LDPC decoding architecture using N parallel processing units with $M = L \times N$.

Memory mapping matrix (B.17) can be reshaped by L as

$$\mathbf{R} = \left[\begin{array}{cccc}
\mathbf{r}_0 & \mathbf{r}_L & \cdots & \mathbf{r}_{(N-1)L} \\
\mathbf{r}_M & \mathbf{r}_{M+L} & \cdots & \mathbf{r}_{M+(N-1)L} \\
\vdots & \vdots & & \vdots \\
\mathbf{r}_{(\alpha-1)M} & \mathbf{r}_{(\alpha-1)M+L} & \cdots & \mathbf{r}_{(\alpha-1)M+(N-1)L} \\
\hline
\mathbf{r}_1 & \mathbf{r}_{L+1} & \cdots & \mathbf{r}_{(N-1)L+1} \\
\mathbf{r}_{M+1} & \mathbf{r}_{M+L+1} & \cdots & \mathbf{r}_{M+(N-1)L+1} \\
\vdots & \vdots & & \vdots \\
\mathbf{r}_{(\alpha-1)M+1} & \mathbf{r}_{(\alpha-1)M+L+1} & \cdots & \mathbf{r}_{(\alpha-1)M+(N-1)L+1} \\
\hline
\vdots & \vdots & & \vdots \\
\hline
\mathbf{r}_{L-1} & \mathbf{r}_{2L-1} & \cdots & \mathbf{r}_{M-1} \\
\mathbf{r}_{M+L-1} & \mathbf{r}_{M+2L-1} & \cdots & \mathbf{r}_{2M-1} \\
\vdots & \vdots & & \vdots \\
\mathbf{r}_{(\alpha-1)M+L-1} & \mathbf{r}_{(\alpha-1)M+2L-1} & \cdots & \mathbf{r}_{\alpha M-1}
\end{array}\right]_{(L\times q\times w_C)\times N} . \quad \text{(B.20)}$$

This convenient message memory mapping allows us to keep the size of the system ROM unchanged, as it only keeps shift and address values for the $C_0^{(c)}$ and $R_0^{(r)}$ groups. For the remaining groups those values can be computed on the fly while decoding. Hence, for $C_\gamma^{(c)}$, shifts are given by

$$\text{shift}_\gamma^{(c)} = (\text{shift}_0^{(c)} + \gamma) \text{ div } L \,, \tag{B.21}$$

and when processing $R_\beta^{(r)}$ addresses are computed as

$$\text{address}_\beta^{(r)} = (\text{address}_0^{(r)} + q \times w_C \times \beta) \mod (q \times w_C \times L) \,, \tag{B.22}$$

where, $\text{shift}_0^{(c)}$ and $\text{address}_0^{(r)}$ are the shift and address values stored in ROM for all $C_0^{(c)}$ and $R_0^{(r)}$ groups.

The shift values stored in ROM are the same for the architectures of both figure 51 and figure 52. Yet because the matrix $\mathbf{R}$ is reshaped the memory addresses contained in $R_0^{(r)}$ change. However, they can be easily obtained by following the procedure described in the previous sub-section, i.e. by searching the matrix (B.20) for the rows where the index r appears.

BIBLIOGRAPHY

[1] *Technical Specification Group Radio Access Network; Base Station (BS) conformance testing (FDD)*, 3GPP Std. 25.141 V9.5.0, Sept. 2010.

[2] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions: with Formulas, Graphs, and Mathematical Tables*, ser. Applied Mathematics Series. National Bureau of Standards, 1972, vol. 5.

[3] A. N. Akansu and R. A. Haddad, *Multiresolution Signal Decomposition*. Academic Press, Inc., 1992.

[4] A. Ambroze, M. Tomlinson, and G. Wade, "Magnitude modulation for small satellite earth terminals using QPSK and OQPSK," in *Proc. IEEE ICC'03*, vol. 3, Anchorage, Alaska, 2003, pp. 2099–2103.

[5] H. Arslan and H. Mahmoud, "Error vector magnitude to SNR conversion for nondata-aided receivers," *IEEE Trans. Wireless Commun.*, vol. 8, no. 5, pp. 2694–2704, May 2009.

[6] C. Berrou, A. Glavieux, and P. Thitimajshima, "Near shannon limit error-correcting coding and decoding: Turbo-codes," in *Proc. IEEE ICC'93*, vol. 2, Geneva, May 1993, pp. 1064–1070.

[7] R. E. Blahut, *Algebraic Codes for Data Transmission*. Cambridge University Press, 2003.

[8] T. Brack, M. Alles, F. Kienle, and N. Wehn, "A Synthesizable IP Core for WIMAX 802.16E LDPC Code Decoding," in *IEEE 17th Int. Symp. on Personal, Indoor and Mobile Radio Commun., PIMRC'06*, Sept. 2006, pp. 1–5.

[9] S. Bregni, *Synchronization of Digital Telecommunications Networks*. John Wiley & Sons, Inc., 2002.

[10] A. B. Carlson, *Communication Systems: An Introduction to Signals and Noise in Electrical Communication, 3rd ed.* McGraw-Hill, 1986.

[11] B. Chatelain and F. Gagnon, "Peak-to-average power ratio and intersymbol interference reduction by nyquist pulse optimization," in *60th Proc. IEEE VTC2004-Fall*, vol. 2, Sept. 2004, pp. 954–958.

[12] J. Chen and M. Fossorier, "Near optimum universal belief propagation based decoding of low-density parity check codes," *IEEE Trans. Commun.*, vol. 50, no. 3, pp. 406–414, March 2002.

[13] M. Chen and O. M. Collins, "Trellis pruning for peak-to-average power ratio reduction," in *Proc. IEEE ISIT2005*, Sept. 2005, pp. 1261–1265.

[14] D. J. Costello and G. D. Forney, "Channel coding: The road to channel capacity," *Proc. IEEE*, vol. 95, no. 6, pp. 1150–1177, June 2007.

[15] R. De Gaudenzi, A. Guillen, and A. Martinez, "Performance analysis of turbo-coded APSK modulations over nonlinear satellite channels," *IEEE Trans. on Wireless Commun.*, vol. 5, no. 9, pp. 2396–2407, Sep. 2006.

[16] ——, "Turbo-coded APSK modulations design for satellite broadband communications," *Int. J. Satellite Commun.*, vol. 24, pp. 261–281, 2006.

[17] J. Dielissen, A. Hekstra, and V. Berg, "Low cost LDPC decoder for DVB-S2," in *Design, Automation and Test in Europe, DATE'06*, vol. 2, March 2006, pp. 1–6.

[18] M. Eroz, F. W. Sun, and L. N. Lee, "DVB-S2 low density parity check codes with near Shannon limit performance," *Int. Journal of Sat. Commun. and Networking*, vol. 22, no. 3, pp. 269–279, May-Jun. 2004.

[19] *Digital Video Broadcasting (DVB); Second generation framing structure, channel coding and modulation systems for Broadcasting, Interactive Services, News Gathering and other broadband satellite applications*, ETSI Std. EN 302 307 V1.1.1, Mar. 2005.

[20] R. G. Gallager, "Low-density parity-check codes," *IRE Trans. Inform. Theory*, vol. IT-8, pp. 21–28, Jan. 1962.

[21] ——, *Low-Density Parity Check Codes*. Cambridge, MA: MIT, 1963.

[22] W. A. Gardner, A. Napolitano, and L. Paura, "Cyclostationarity: Half a century of research," *Signal Processing (Elsevier)*, vol. 86, no. 4, pp. 639–697, Apr. 2006.

[23] F. Garvan, *The Maple Book*. Chapman & Hall/CRC, 2002.

[24] A. Georgiadis, "Gain, phase imbalance, and phase noise effects on error vector magnitude," *IEEE Trans. Veh. Technol.*, vol. 53, no. 2, pp. 443–449, Mar. 2004.

[25] G. B. Giannakis, "Cyclostationary signal analysis," in *The Digital Signal Processing Handbook*, V. K. Madisetti and D. B. Williams, Eds. Chapman and Hall/CRCnetBase, 1999, ch. 17.

[26] W. J. Gilbert and W. K. Nicholson, *Modern Algebra with Applications, 2nd ed.* John Wiley & Sons, 2004.

[27] M. J. E. Golay, "Notes on digital coding," *Proc. IRE*, vol. 37, p. 657, July 1974.

[28] M. Gomes, G. Falcão, J. Goncalves, V. Silva, M. Falcao, and P. Faia, "HDL library of processing units for generic and DVB-S2 LDPC decoding," in *Int. Conf. on Signal Proc. and Multimedia Applications, SIGMAP'06*, 2006, pp. 17–24.

[29] M. Gomes, V. Silva, C. Neves, and R. Marques, "Serial ldpc decoding on a SIMD DSP using horizontal scheduling," in *14th European Signal Processing Conference, EUSIPCO'06*, Florence, Italy, September 2006.

[30] M. Gomes, G. Falcão, A. Sengo, V. Ferreira, V. Silva, and M. Falcão, "High throughput encoder architecture for DVB-S2 LDPC-IRA codes," in *Proc. IEEE Int. Conf. Microelectronics, ICM'07*, Cairo, Egypt, Dec. 2007, pp. 271–274.

[31] M. Gomes, G. Falcão, V. Silva, V. Ferreira, A. Sengo, and M. Falcão, "Flexible parallel architecture for DVB-S2 LDPC decoders," in *Proc. IEEE Globecom'07*, Washington, DC, Nov. 2007, pp. 3265–3269.

[32] M. Gomes, F. Cercas, V. Silva, and M. Tomlinson, "Efficient m-QAM transmission using compacted magnitude modulation tables," in *Proc. IEEE Globecom'08*, New Orleans, LA, Nov. 2008.

[33] M. Gomes, G. Falcão, V. Silva, V. Ferreira, A. Sengo, L. Silva, N. Marques, and M. Falcão, "Scalable and parallel codec architectures for the DVB-S2 FEC system," in *Proc. IEEE APCCAS'08*, Macau, China, Nov. 2008, pp. 1506–1509.

[34] M. Gomes, V. Silva, F. Cercas, and M. Tomlinson, "Low back-off 16-APSK transmission using magnitude modulation and symbol quantization," in *IEEE Int. Workshop on Satellite and Space Commun., IWSSC'08*, Toulouse, France, Oct. 2008, pp. 229–233.

[35] M. Gomes, F. Cercas, V. Silva, and M. Tomlinson, "Polyphase magnitude modulation for peak power control," in *17th European Signal Processing Conference, EUSIPCO 2009*, Glasgow, Scotland, August 2009.

[36] M. Gomes, V. Silva, F. Cercas, and M. Tomlinson, "Power efficient back-off reduction through polyphase filtering magnitude modulation," *IEEE Commun. Lett.*, vol. 13, no. 8, pp. 606–608, August 2009.

[37] ——, "Real-time LUT-less magnitude modulation for peak power control of single carrier rrc filtered signals," in *IEEE 10th Workshop on Signal Proc. Adv. in Wireless Commun., SPAWC'09*, Perugia, Italy, June 2009, pp. 424–428.

[38] M. Gomes, F. Cercas, V. Silva, and M. Tomlinson, "Magnitude modulation for VSAT's low back-off transmission," *Journal Commun. Networks (JCN), special issue on Recent Adv. in Satell. and Space Commun.*, vol. 12, no. 6, pp. 544–557, Dec. 2010.

[39] M. Gomes, V. Silva, F. Cercas, and M. Tomlinson, "Analytical analysis of polyphase magnitude modulation method's performance," in *Proc. IEEE ICC'2010*, Cape Town, South Africa, May 2010.

[40] M. Hahm, "Device and method for limiting peaks of a signal," U.S. Patent 6 356 606 B1, 2002.

[41] R. W. Hamming, "Error detecting and error correcting codes," *Bell Syst. Tech. J.*, vol. 29, no. 2, pp. 147–160, April 1950.

[42] S. H. Han and J. H. Lee, "An overview of peak-to-average power ratio reduction techniques for multicarrier transmission," *IEEE Wireless Commun. Mag.*, vol. 12, no. 2, pp. 56–65, April 2005.

[43] L. Hanzo, W. T. Webb, and T. Keller, *Single- and Multi-carrier Quadrature Amplitude Modulation: Principles and Applications for Personal Communications, WLANs and Broadcasting*. Wiley, 2000.

[44] F. Harris, *Multirate Signal Processing for Communication Systems*. Prentice Hall PTR, 2004.

[45] M. H. Hayes, *Statistical Digital Signal Processing and Modeling*. John Wiley & Sons, Inc., 1992.

[46] H. Holma and A. Toskala, Eds., *LTE for UMTS-OFDMA and SC-FDMA based radio access*. John Wiley & Sons Ltd, 2009.

[47] X.-Y. Hu, E. Eleftheriou, D.-M. Arnold, and A. Dholakia, "Efficient implementations of the sum-product algorithm for decoding LDPC codes," in *Proc. IEEE Globecom'01*, vol. 2, San Antonio, Texas, Nov. 2001, pp. 1036–1036.

[48] M. Hunton, "System and method for post filtering peak power reduction in communications systems," U.S. Patent 7 170 952 B2, 2007.

[49] *Part 16: Air Interface for Fixed and Mobile Broadband Wireless Access Systems - Amendment for Physical and Medium Access Control Layers for Combined Fixed and Mobile Operation in Licensed Bands*, IEEE Std. 802.16e, 2005.

[50] *IEEE Standard for Local and Metropolitan Area Networks Part 20: Air Interface for Mobile Broadband Wireless Access Systems Supporting Vehicular Mobility - Physical and Media Access Control Layer Specification*, IEEE Std. 802.20-2008, 2008.

[51] *IEEE Standard for Information technology-Telecommunications and information exchange between systems-Local and metropolitan area networks-Specific requirements Part 3: Carrier Sense Multiple Access with Collision Detection (CSMA/CD) Access Method and Physical Layer Specifications*, IEEE Std. 802.3an-2006, 2006.

[52] K. A. S. Immink, "Reed-solomon codes and the compact disc," in *Reed-Solomon Codes and Their Applications*, S. B. Wicker and V. K. Bhargava, Eds. NJ: IEEE Press, 1994, pp. 41–59.

[53] L. J. Ippolito Jr., *Satellite communications systems engineering : atmospheric effects, satellite link design, and system performance*, ser. Wiley series on wireless communications and mobile computing. John Wiley & Sons Ltd, 2008.

[54] ITU, *Handbook on satellite communications, 3rd ed.* John Wiley & Sons Ltd, 2002.

[55] B. Jelonnek, "Method and system for clipping a baseband input signal," U.S. Patent 0 036 251 A1, 2007.

[56] ——, "Method for limiting amplitude of a carrier frequency output signal," U.S. Patent 7 327 804 B2, 2008.

[57] T. Jiang and Y. Wu, "An overview: Peak-to-average power ratio reduction techniques for OFDM signals," *IEEE Trans. Broadcasting*, vol. 54, no. 2, pp. 257–268, Jun. 2008.

[58] H. Jin, A. Khandekar, and R. McEliece, "Irregular repeat-accumulate codes," in *Proc. 2nd Int. Symp. on Turbo Codes & Related Topics*, Brest, France, 2000, pp. 1–8.

[59] P. B. Kenington, *High-linearity RF Amplifier Design*. Artech House, 2000.

[60] F. Kienle, T. Brack, and N. Wehn, "A synthesizable IP core for DVB-S2 LDPC code decoding," in *Design, Automation and Test in Europe, DATE'05*, vol. 3, March 2005, pp. 100–105.

[61] P. Koopman and T. Chakravarty, "Cyclic redundancy code (CRC) polynomial selection for embedded networks," in *Proc. Int. Conf. on Dependable Systems and Networks (DSN'04)*, Florence, Italy, June 2000, pp. 145–154.

[62] Y. Kou, S. Lin, and M. Fossorier, "Low-density parity-check codes based on finite geometries: a rediscovery and new results," *IEEE Trans. Inf. Theory*, vol. 47, no. 7, pp. 2711–2736, Nov. 2001.

[63] N. Lashkarian, E. Hemphill, H. Tarn, H. Parekh, and C. Dick, "Reconfigurable digital front-end hardware for wireless base-station transmitters: Analysis, design and fpga implementation," *IEEE Trans. Circuits Syst. I: Regular Papers*, vol. 54, no. 8, pp. 1666–1677, Aug. 2007.

[64] V. K. N. Lau, "Average of peak-to-average ratio (PAR) of IS95 and CDMA2000 systems-single carrier," *IEEE Commun. Lett.*, vol. 5, no. 4, pp. 160–162, Apr. 2001.

[65] E. A. Lee and D. G. Messerschmitt, *Digital Communication*. Kluwer Academic Publishers, 1990.

[66] S. Lin and J. D. J. Costello, *Error Control Coding, 2nd ed.* NJ: Pearson Prentice Hall, 2004.

[67] S. Litsyn, *Peak Power Control in Multicarrier Communications*. Cambridge University Press, 2007.

[68] D. J. C. MacKay and R. M. Neal, "Good codes on very sparse matrices," in *Proc. Cryptography Coding, 5th IMA Conf.*, E. C. Boyd, Ed. Berlin, Germany: Springer, 1995, pp. 100–111.

[69] ——, "Near shannon limit performance of low density parity check codes," *Electron. Letters*, vol. 32, no. 18, pp. 1645–1646, Aug. 1996.

[70] G. Maral, *Satellite communications systems, 5th ed.* John Wiley & Sons Ltd, 2009.

[71] MATLAB, *version 7.9.0.529 (R2009b)*. Natick, Massachusetts: The MathWorks Inc., 2009.

[72] U. Mengali and A. N. D'Andrea, *Synchronization Techniques for Digital Receivers*. Plenum Press, 1997.

[73] H. Meyr, M. Moeneclaey, and S. Fechtel, *Digital Communication Receivers: Synchronization, Channel Estimation, and Signal Processing*. John Wiley & Sons, Inc., 1998.

[74] S. Miller and R. O'Dea, "Peak power and bandwidth efficient linear modulation," *IEEE Trans. Commun.*, vol. 46, no. 12, pp. 1639–1648, Dec. 1998.

[75] ——, "Radio with peak power and bandwidth efficient modulation," U.S. Patent 5 621 762, 1997.

[76] S. K. Mitra, *Digital Signal Processing: A Computer Based Approach - 3th ed.* McGraw-Hill, 2006.

[77] T. K. Moon, *Error Correction Coding: Mathematical Methods and Algorithms*. John Wiley & Sons, 2005.

[78] R. H. Morelos-Zaragoza, *The Art of Error Correcting Coding, 2nd ed.* NY: John Wiley & Sons, 2006.

[79] B. Moris, "Peak power reduction using windowing and filtering," U.S. Patent 7 013 161 B2, 2006.

[80] J. Moura, J. Lu, and H. Zhang, "Structured low-density parity-check codes," *IEEE Signal Process. Mag.*, vol. 21, no. 1, pp. 42–55, Jan. 2004.

[81] O'Dea, "Radio with peak power and bandwidth efficient modulation using asymmetric symbol constellations," U.S. Patent 5 606 578, 1997.

[82] A. V. Oppenheim, *Signals & Systems (2nd ed.)*. Prentice-Hall Signal Processing Series, 1996.

[83] A. V. Oppenheim, R. W. Schaffer, and J. R. Buck, *Discrete-Time Signal Processing – 2nd ed.* Prentice-Hall Signal Processing Series, 1998.

[84] A. Papoulis and S. U. Pillai, *Probability, Random Variables and Stochastic Processes (4th ed.)*. McGraw-Hill, 2002.

[85] M. Peeters, "Power reduction," U.S. Patent 6 931 079 B1, 2005.

[86] J. Pinto and I. Darwazeh, "Error vector magnitude relation to magnitude and phase distortion in 8-PSK systems," *Electronics Letters*, vol. 37, no. 7, pp. 437–438, mar. 2001.

[87] J. G. Proakis, *Digital Communications - 4th ed.* McGraw-Hill, 2000.

[88] P. Rha and S. Hsu, "Peak-to-average ratio (PAR) reduction by pulse shaping using a new family of generalized raised cosine filters," in *Proc. IEEE VTC2003-Fall*, vol. 1, Oct. 2003, pp. 706–710.

[89] T. Richardson and R. Urbanke, *Modern Coding Theory*. Cambridge University Press, 2008.

[90] T. Richardson, M. Shokrollahi, and R. Urbanke, "Design of capacity-approaching irregular low-density parity-check codes," *IEEE Trans. Inf. Theory*, vol. 47, no. 2, pp. 619–637, Feb. 2001.

[91] W. E. Ryan, "An Introduction to LDPC Codes," in *Advanced error control techniques for data storage systems*, E. Kurtas and B. Vasic, Eds. CRC Press, 2006, ch. 12.

[92] A. Sandri and A. Spalvieri, "Method and circuits for reducing the peak power of the filtered signal transmitted in a digital link," U.S. Patent 5 493 587, 1996.

[93] V. Sathe and P. Vaidyanathan, "Effects of multirate systems on the statistical properties of random signals," *IEEE Trans. Signal Process.*, vol. 41, no. 1, pp. 131–146, Jan. 1993.

[94] A. S. Sedra and K. C. Smith, *Microelectronic Circuits, 4th ed.* Oxford University Press, 1998.

[95] S. Sesia, M. Baker, and I. Toufik, *LTEÐthe UMTS long term evolution : from theory to practice*. John Wiley & Sons Ltd, 2009.

[96] J. S. Seybold, "Output back-off requirements for root-raised, cosine-filtered digital signals," *RF Signal Processing*, pp. 50–58, Jun. 2002.

[97] Seyed-Esfahani, "Method and system to reduce PAPR in single carrier and SCBT communications systems," U.S. Patent 0 239 038 A1, 2010.

[98] R. A. Shafik, S. Rahman, and A. Razibul Islam, "On the extended relationships among EVM, BER and SNR as performance metrics," in *Int. Conf. on Electr. and Comp. Engin., ICECE'06*, Dec. 2006, pp. 408–411.

[99] C. E. Shannon, "A mathematical theory of communication," *Bell Syst. Tech. J.*, vol. 27, pp. 379–423 and 623–656, 1948.

[100] E. Sharon, S. Litsyn, and J. Goldberger, "An efficient message-passing schedule for LDPC decoding," in *Electrical and Electronics Engineers in Israel, 2004. Proceedings. 2004 23rd IEEE Convention of*, Sept. 2004, pp. 223–226.

[101] B. Sklar and F. Harris, "The ABCs of linear block codes," *IEEE Signal Process. Mag.*, vol. 21, no. 4, pp. 14–35, July 2004.

[102] S. Takabayashi, "Peak power suppressing apparatus and peak power suppressing method," U.S. Patent 0 031 380 A1, 2008.

[103] M. Tanahashi and H. Ochiai, "A simple modulation code with peak power reduction and coding gain," in *Proc. IEEE ICC'08*, May 2008, pp. 1365–1369.

[104] ——, "Near constant envelope trellis shaping for psk signaling," *IEEE Trans. Commun.*, vol. 57, no. 2, pp. 450–458, Feb. 2009.

[105] R. Tanner and J. Woodard, *WCDMA – Requirements and Practical Design*. John Wiley & Sons Ltd, 2004.

[106] R. M. Tanner, "A recursive approach to low complexity codes," *IEEE Trans. Inf. Theory*, vol. IT-27, no. 5, pp. 533–547, Sept. 1981.

[107] C. W. Therrien, "Issues in multirate statistical signal processing," in *35th Asilomar Conf. on Signals, Systems and Computers*, vol. 1, 2001, pp. 573–576.

[108] A. Tomlinson, A. Ambroze, and G. Wade, "Power and bandwidth efficient modulation and coding for small satellite communication terminals," in *Proc. IEEE ICC'02*, vol. 5, New York, April 2002, pp. 2943–2946.

[109] O. Väänänen, "Digital modulators with crest factor reduction techniques," Ph.D. dissertation, Helsinki University of Technology, Finland, 2006.

[110] M. Wang, W. Xiao, and T. Brown, "Soft decision metric generation for QAM with channel estimation error," *IEEE Trans. Commun.*, vol. 50, no. 7, pp. 1058–1061, Jul. 2002.

[111] D. Wulich and L. Goldfeld, "Bound of the distribution of instantaneous power in single carrier modulation," *IEEE Trans. Wireless Commun.*, vol. 4, no. 4, pp. 1773–1778, Jul. 2005.

[112] H. Xiao and A. Banihashemi, "Graph-based message-passing schedules for decoding LDPC codes," *IEEE Trans. Commun.*, vol. 52, no. 12, pp. 2098–2105, Dec. 2004.

[113] C.-Y. Yao, "The design of square-root-raised-cosine FIR filters by an iterative technique," *IEICE Trans. Fundam. Electron. Commun. Comput. Sci.*, vol. E90-A, pp. 241–248, Jan. 2007.