UNIVERSITÀ DEGLI STUDI DI PARMA

DIPARTIMENTO DI INGEGNERIA DELL'INFORMAZIONE

Dottorato di Ricerca in Tecnologie dell'Informazione
XXVI Ciclo

Andrea Modenini

Advanced transceivers for spectrally-efficient communications

Dissertazione presentata per il conseguimento
del titolo di Dottore di Ricerca

Gennaio 2014

UNIVERSITÀ DEGLI STUDI DI PARMA

Dottorato di Ricerca in Tecnologie dell'Informazione

XXVI Ciclo

Advanced transceivers for spectrally-efficient communications

Coordinatore:

Chiar.mo Prof. Marco Locatelli

Tutor:

Chiar.mo Prof. Giulio Colavolpe

Dottorando: *Andrea Modenini*

Gennaio 2014

to my adorable and wonderful wife,
to my beloved family,
to my dear friends,
and all those that supported me through time and believed in me
(e.g. Giulio)

Contents

List of Figures

List of Tables

Introduction

Vanity, definitely my favorite sin.

– The Devil's Advocate

TELECOMMUNICATIONS are a growing field in the global industry, and their request is increasing more and more every year. Due to this growing demand, the available bandwidths are getting insufficient to fulfill the global request. Independently from the kind of communication system (satellite, wireless, optical), the desire of every telecommunications operator is to transmit at the highest possible rate in the available bandwidth for a given power. In more technical words, the aim is to maximize the *spectral efficiency* of the communication systems.

In this thesis, we will consider techniques to improve the spectral efficiency of digital communication systems, operating on the whole transceiver scheme. First, we will focus on receiver schemes having detection algorithms with a complexity constraint. We will optimize the parameters of the reduced detector with the aim of maximizing the *achievable information rate*. Namely, we will adopt the *channel shortening* technique (see [1, 2] and references therein).

Then, we will focus on a technique that is getting very popular in the last years (although presented for the first time in 1975): *faster-than-Nyquist* signaling, and its extension which is *time packing* (see [3, 4, 5] and references therein). Time packing is a very simple technique that consists in introducing intersymbol interference on purpose with the aim of increasing the spectral efficiency of finite order constellations.

Finally, in the last chapters we will combine all the presented techniques, and we will consider their application to satellite channels.

Although we will not consider here optical communications, we point out that many of these techniques, can be applied (with suitable *tweaks*) also to these scenarios. It is worth to cite [6, 7, 8].

The remainder of this thesis is organized as follows: Chapter 1 will introduce the basics of the work in this thesis. Chapter 2 will focus entirely on the channel shortening technique, illustrating also many practical detection schemes. In Chapter 3 we will turn our attention to the time packing technique, showing its potential. Finally in Chapters 4 and 5 we will *connect the dots* and apply the proposed techniques to the satellite channel.

Publications

The work in this thesis is part of the results of the research activities conducted during the PhD studies, with the following publications:

Journals

- A. Modenini, F. Rusek, and G. Colavolpe "Optimal transmit filters for ISI channels under channel shortening detection," IEEE Transactions on Communications, vol. 61, pp. 4997-5005, December 2013.

- A. Piemontese, A. Modenini, G. Colavolpe, and N. Alagha "Improving the spectral efficiency of nonlinear satellite systems through time-frequency packing and advanced processing," IEEE Transactions on Communications, vol. 61, pp. 3404-3412, August 2013.

- G. Colavolpe, A. Modenini, and F. Rusek, "Channel Shortening for Nonlinear Satellite Channels," Communications Letters, IEEE , vol.16, no.12, pp.1929-1932, December 2012.

- G. Colavolpe, Tommaso Foggi, A. Modenini, and A. Piemontese, "Faster-than-Nyquist and beyond: how to improve spectral efficiency by accepting interference," Opt. Express 19, 26600-26609 (2011).

Conferences

- A. Piemontese, A. Modenini, G. Colavolpe, and N. Alagha, "Spectral Efficiency of Time-Frequency-Packed Nonlinear Satellite Systems," in 31th AAIA International communications satellite systems conference, Florence, Italy, October 2013.

- G. Colavolpe, and A. Modenini, "Iterative carrier syncrhonization in the absence of distributed pilots for low SNR applications," in Proc. Intern. Workshop of Tracking Telemetetry and Command System for Space Communications. (TTC'13), European Space Agency, Darmstadt, Germany, September 2013.

- A. Modenini, F. Rusek, and G. Colavolpe, "Optimal transmit filters for constrained complexity channel shortening detectors," in Proc. IEEE Intern. Conf. Commun. (ICC'13), Budapest, Hungary, June 2013, pp. 1688-1693.

- A. Modenini, G. Colavolpe, and N. Alagha, "How to significantly improve the spectral efficiency of linear modulations through time-frequency packing and advanced processing," in Proc. IEEE Intern. Conf. Commun. (ICC'12), Ottawa, Canada, June 2012, pp. 3430-3434.

Patents

- G. Colavolpe, A. Modenini, A. Piemontese, and N. Alagha, "Data detection method and data detector for signals transmitted over a communication channel with inter-symbol interference," assigned to ESA-ESTEC, The Neederlands. International patent application n. F027800186/WO/PCT, December 2012.

Common abbreviations

AWGN	additive white Gaussian noise
AIR	achievable information rate
ASE	achievable spectral efficiency
BCJR	Bahl Cocke Jelinek Raviv (algorithm)
CS	channel shortening
DTFT	discrete time Fourier transform
DFT	discrete Fourier transform
HPA	high power amplifier
ICI	interchannel interference
IMUX	input multiplexer (filter)
IR	information rate
ISI	intersymbol interference
FTN	faster-than-Nyquist
MAP	maximum a posteriori
MF	matched filter
MIMO	multiple-input multiple-output
MMSE	minimum mean square error
OMUX	output multiplexer (filter)
PSK	phase shift keying (modulation)
QAM	quadrature amplitude (modulation)
SE	spectral efficiency
WF	whitening filter
WMF	whitening matched filter

Mathematical notation

h	scalar (possibly complex)
h^*	complex-conjugated of the complex scalar

$\Re(h)$	real part of of the complex scalar
$\Im(h)$	imaginary part of the complex scalar
$\boldsymbol{h}$	vector
$\boldsymbol{h}^T$	transposed vector
$\boldsymbol{h}^\dagger$	transposed and conjugated vector (Hermitian)
$\boldsymbol{H}$	matrix
$\boldsymbol{I}$	identity matrix
$(\boldsymbol{H})_{ij}$	scalar entry (i,j) of the matrix
$\boldsymbol{H}^T$	transposed matrix
$\boldsymbol{H}^\dagger$	transposed and conjugated matrix (Hermitian)
$\mathrm{Tr}(\boldsymbol{H})$	trace of the matrix
$\mathbf{h}$	block vector
$\mathbf{h}^T$	transposed block vector
$\mathbf{h}^\dagger$	transposed and conjugated block vector (Hermitian)
$\mathbf{H}$	block matrix
$(\mathbf{H})_{ij}$	matrix entry (i,j) of the block matrix
$\mathbf{H}^T$	transposed block matrix
$\mathbf{H}^\dagger$	transposed and conjugated block matrix (Hermitian)
δ_i	Kronecker delta
$\otimes$	convolution
$F[y(\omega)]$	functional of $y(\omega)$
$\frac{\delta F[y(\omega)]}{\delta y}$	functional derivative w.r.t. $y(\omega)$
$P(c)$	probability mass function of a discrete random variable c
$H(c)$	entropy of a discrete random variable c
$p(r)$	probability density function of a continuous random variable r
$h(r)$	entropy of a continous random variable r

Chapter 1

Basics

THIS chapter will introduce the basic arguments on which we will mainly focus in this thesis. The chapter is organized as follows: in §1.1 we introduce the notation for linear modulations. In §1.2 we describe the main observation models. In §1.3 we describe optimal detection algorithms for the presented observation models. Finally, in §1.4 we present the figures of merit that will be used for the performance analysis.

1.1 Linear modulations over the AWGN channel

In this thesis, we will mainly consider linearly modulated signals whose complex envelope can be expressed as

$$s(t) = \sum_{k=0}^{N-1} c_k p(t-kT) \tag{1.1}$$

being $\{c_k\}_{k=0}^{N-1}$ the transmitted symbols, $p(t)$ the shaping pulse, and T the symbol time. Symbols $\{c_k\}$ will be considered belonging to a M-ary constellation in the complex domain. The transmitted symbols can be either coded or uncoded. If the signal is transmitted over a channel with additive white Gaussian noise (AWGN), the complex envelope of the received signal will read

$$\begin{aligned} r(t) &= s(t)+w(t) && (1.2)\\ &= \sum_{k=0}^{N-1} c_k p(t-kT)+w(t) && (1.3) \end{aligned}$$

where $w(t)$ is white Gaussian noise having power spectral density N_0. Without loss of generality, in (1.3) we considered a frequency flat channel. Clearly, the extension to frequency selective channels is obtained straightforwardly: the received shaping pulse in (1.3) will be equal to

$$p(t)\otimes h(t) \qquad (1.4)$$

being $h(t)$ the channel impulse response.

1.2 Observation models

For detection, we need a discrete-time observation model $\boldsymbol{r}$ of the received signal. The observation model $\boldsymbol{r}$ shall be a sufficient statistics, i.e., a function of the received signal that does not involve any information loss. For the received signal (1.3), different sufficient statistics can be found.

The first model that we consider is the *Ungerboeck observation model* [9]. The Ungerboeck observation model is obtained as shown in Figure 1.1. The received signal passes through a filter matched to the shaping pulse $p(t)$ (matched filter, MF). The signal at the output is then sampled with time interval T. The sequence of samples $\boldsymbol{r}=[r_0,\ldots,r_{N-1}]^T$ results to be

$$r_k=\sum_{i=-\nu}^{\nu} c_{k-i}g_i+n_k \qquad (1.5)$$

where

$$g_i=\int_{-\infty}^{\infty} p(t)p^*(t-iT)\mathrm{d}t \qquad (1.6)$$

are the intersymbol interference (ISI) taps, which are null for $|i|>\nu$ (being ν the channel memory), and $\{n_k\}$ are Gaussian random variables with autocorrelation function

$$\mathrm{E}\{n_{k+i}n_k^*\}=N_0g_i\,. \qquad (1.7)$$

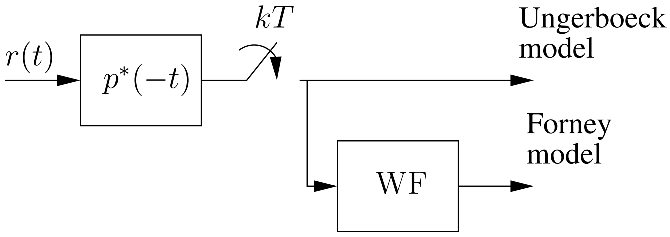

Figure 1.1: Block diagram of the system which carry out the Ungerboeck observation model and the Forney observation model.

The samples (1.5) can be gathered in a useful matrix notation

$$\boldsymbol{r} = \boldsymbol{G}\boldsymbol{c} + \boldsymbol{n} \tag{1.8}$$

where $\boldsymbol{c}$ and $\boldsymbol{n}$ are defined as $\boldsymbol{c} = [c_0, \ldots, c_{N-1}]^T$, $\boldsymbol{n} = [n_0, \ldots, n_{N-1}]^T$, and $\boldsymbol{G}$ is a Toeplitz matrix (see Appendix A) with entries $(\boldsymbol{G})_{\ell m} = g_{\ell-m}$.

Since white noise is often preferred, the MF output can be filtered by a whitening filter (WF)[1] as shown in Figure 1.1. This yields another sufficient statistics known as *Forney observation model* [10]. The Forney model reads

$$r_k = \sum_{i=0}^{\nu} c_{k-i} h_i + w_k \tag{1.9}$$

where $\{w_k\}$ are Gaussian random variables with $\mathrm{E}\{w_{k+i} w_k\} = N_0 \delta_i$ being δ_i the Kronecker delta, and $\{h_i\}_{i=0}^{\nu}$ are the ISI taps such that $g_i = h_i \otimes h^*_{-i}$. The value of memory ν of the Forney model is always equal to the one of the Ungerboeck model. The Forney observation model (1.9) can be expressed by means of the matrix notation

$$\boldsymbol{r} = \boldsymbol{H}\boldsymbol{c} + \boldsymbol{w} \tag{1.10}$$

where $\boldsymbol{H}$ is a Toeplitz and lower triangular matrix with entries $(\boldsymbol{H})_{\ell m} = h_{\ell-m}$. Moreover it holds $\boldsymbol{H}^\dagger \boldsymbol{H} = \boldsymbol{G}$.

[1] The cascade of the MF and the WF is called whitened matched filter (WMF).

Both the Ungerboeck and Forney models show that a discrete-time channel equivalent to the continuous-time one can be found. In addition to the Ungerboeck and Forney models, there exist other sets of sufficient statistics (although not consider in this thesis). Another example of sufficient statistics is the one described in [11].

1.3 Optimal MAP detection: the BCJR algorithm

Optimal maximum a posteriori (MAP) symbol detection is based on the strategy

$$\hat{c}_k = \arg\max_{c_k} P(c_k|\boldsymbol{r}) \qquad k = 0, \ldots, N-1\,. \tag{1.11}$$

The *a posteriori probabilities* $P(c_k|\boldsymbol{r})$ can be effectively computed by means of the Bahl-Cocke-Jelinek-Raviv (BCJR) algorithm [12]. Let us define σ_k as the *state* of the channel at the discrete time k, and gather all the states in the vector $\boldsymbol{\sigma} = [\sigma_0, \ldots, \sigma_{N-1}]^T$. The basic hypothesis of the BCJR algorithm is that it exists a definition of the state σ_k which allows to factorize the probability $P(\boldsymbol{c}, \boldsymbol{\sigma}|\boldsymbol{r})$ as

$$P(\boldsymbol{c}, \boldsymbol{\sigma}|\boldsymbol{r}) \propto \prod_{k=0}^{N-1} \Lambda_k(r_k, c_k, \sigma_k, \sigma_{k+1}) P(c_k) \tag{1.12}$$

where $P(c_k)$ is the *a priori probability* on the symbol c_k at time k, and $\Lambda_k(r_k, c_k, \sigma_k, \sigma_{k+1})$ is the *metric* at time k. The metrics are not necessarily probability mass functions.

The BCJR algorithm can be derived as follow. Equation (1.12) can be represented by the factor graph (FG, see [13]) in Figure 1.2. The application of the sum product algorithm (SPA) to the FG in Figure 1.2 gives a unique message passing algorithm to compute any marginal probability of (1.12). In particular, we are interested in the a posteriori probabilities in (1.11) that are obtained with the following marginalization:

$$P(c_k|\boldsymbol{r}) = \sum_{\sim\{c_k\}} P(\boldsymbol{c}, \boldsymbol{\sigma}|\boldsymbol{r})\,, \tag{1.13}$$

where $\sum_{\sim\{c_k\}}$ denotes the sum with respect to all the variables, except c_k. Denoting the messages on the graph by $\alpha_k(\sigma_k)$ and $\beta_k(\sigma_k)$ as shown in in Figure 1.3, it can be shown that the a posteriori probabilities are obtained with the following message passing algorithm:

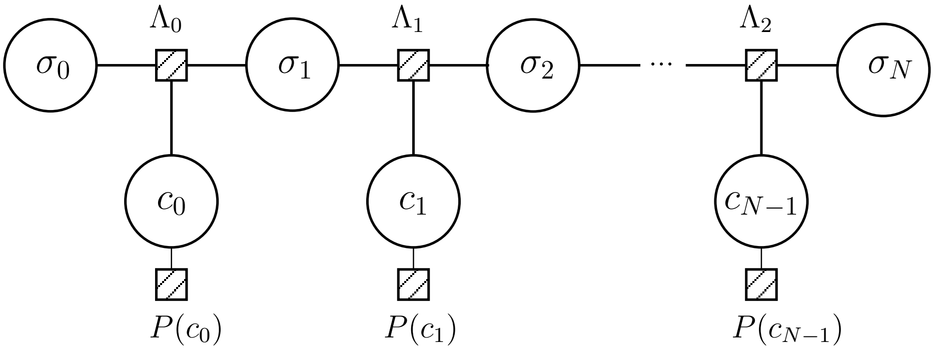

Figure 1.2: Factor graph of the BCJR algorithm.

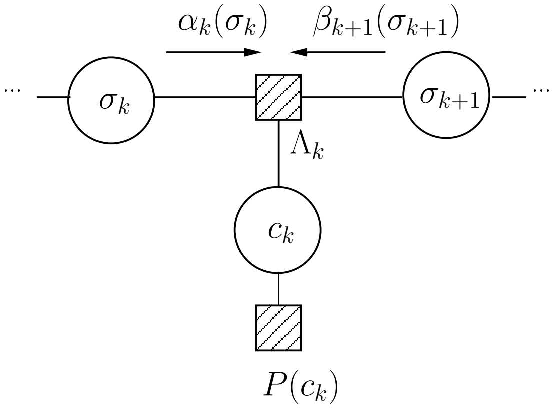

Figure 1.3: Message passing algorithm on a section of the BCJR factor graph.

- Initialize the algorithm as

$$\begin{aligned}\alpha_0(\sigma_0) &= 1 &(1.14)\\ \beta_N(\sigma_N) &= 1\,. &(1.15)\end{aligned}$$

- *Forward recursion*: for each $k = 0, \ldots, N-1$ compute the messages $\alpha_{k+1}(\sigma_{k+1})$ as

$$\alpha_{k+1}(\sigma_{k+1}) = \sum_{c_k,\sigma_k} \alpha_k(\sigma_k)\Lambda_k(r_k, c_k, \sigma_k, \sigma_{k+1})P(c_k)\,. \tag{1.16}$$

- *Backward recursion*: for each $k = N-1, N-2, \ldots, 0$ compute the messages $\beta_k(\sigma_k)$ as

$$\beta_k(\sigma_k) = \sum_{c_k,\sigma_{k+1}} \beta_{k+1}(\sigma_{k+1})\Lambda_k(r_k, c_k, \sigma_k, \sigma_{k+1})P(c_k)\,. \tag{1.17}$$

- *Completion*: for each $k = 0, \ldots, N-1$ compute the a posteriori probabilities as

$$P(c_k|\mathrm{r}) \propto P(c_k) \sum_{\sigma_k,\sigma_{k+1}} \alpha(\sigma_k)\Lambda_k(r_k, c_k, \sigma_k, \sigma_{k+1})\beta(\sigma_{k+1})\,. \tag{1.18}$$

The complexity of the BCJR algorithm is proportional to $\mathscr{O}(MS)$, being M the cardinality of the transmitted symbols $\{c_k\}$ and S the cardinality of the state $\{\sigma_k\}$. The algorithm is conveniently implemented in the logarithmic domain [14]. The reader can notice that the BCJR algorithm performs a trellis processing similar to the Viterbi algorithm. In fact the algorithms are stricly related. It can be shown that working in the logarithm domain, and by substituting the logarithm of a sum of exponentials with the max of the arguments, the Viterbi algorithm is obtained [13].

For the sake of clarity we finally show the BCJR algorithm for the observation models presented in §1.2. For the Forney model, it is easy to notice from (1.9) that the state can be defined by the vector of the past symbols $\boldsymbol{\sigma}_k = [c_{k-1}, \ldots, c_{k-\nu}]$. Thus, the conditional probability factorizes with terms

$$\begin{aligned}\Lambda_k(c_k, \boldsymbol{\sigma}_k, \boldsymbol{\sigma}_{k+1}) &= p(r_k|c_k, \boldsymbol{\sigma}_k)\mathscr{I}(c_k, \boldsymbol{\sigma}_k, \boldsymbol{\sigma}_{k+1}) &(1.19)\\ &= \exp\left\{-\frac{|y_k - \sum_{i=0}^{\nu} h_i c_{k-i}|^2}{N_0}\right\}\mathscr{I}(c_k, \boldsymbol{\sigma}_k, \boldsymbol{\sigma}_{k+1}) &(1.20)\end{aligned}$$

where $\mathscr{I}(c_k, \boldsymbol{\sigma}_k, \boldsymbol{\sigma}_{k+1})$ is an indicator function: it is equal to 1 if the transition $(c_k, \sigma_k) \to \sigma_{k+1}$ is valid, and to 0 otherwise.

Since the cardinality of the state is $S = M^\nu$, the complexity of the algorithm increases exponentially with ν. Moreover it can be shown that for the Forney model the forward and backward recursions have the following probabilistic meaning [15]

$$\begin{aligned} \alpha_k(\sigma_k) &= P(\sigma_k | \boldsymbol{r}_0^{k-1}) \quad \forall k = 0, \ldots, N && (1.21) \\ \beta_k(\sigma_k) &= p(\boldsymbol{r}_k^{N-1} | \sigma_k) \quad \forall k = 0, \ldots, N && (1.22) \end{aligned}$$

where $\boldsymbol{r}_a^b$ denotes either the vector $[r_a, \ldots, r_b]$ for any $a \leq b$, or the empty set otherwise.

For the Ungerboeck observation model it can be shown (see [16]) that the state can be defined again as $\boldsymbol{\sigma}_k = [c_{k-1}, \ldots, c_{k-\nu}]$ and the conditional probability factorizes with terms

$$\Lambda_k(c_k, \boldsymbol{\sigma}_k, \boldsymbol{\sigma}_{k+1}) = \exp\left\{ \frac{2\Re\left(c_k^* r_k\right) - |c_k|^2 g_0 - 2c_k^* \sum_{i=1}^{\nu} g_i c_{k-i}}{N_0} \right\} \mathscr{I}(c_k, \boldsymbol{\sigma}_k, \boldsymbol{\sigma}_{k+1}) . \tag{1.23}$$

Since the memory ν of the Ungerboeck model in (1.5) is always equal to the memory of the Forney model, also the complexity of the algorithms is the same. We point out that for the Ungerboeck model the $\Lambda_k(c_k, \boldsymbol{\sigma}_k, \boldsymbol{\sigma}_{k+1})$ are not probability density functions as for the Forney case.

1.4 Performance analysis

We will consider different figures of merit for the performance analysis. The first figure of merit is the *spectral efficiency* (SE) that can be achieved by a given modulation and coding format (MODCOD) on a given channel defined as

$$\mathrm{SE} = \frac{r \log_2(M)}{TW} \quad [\mathrm{bit/s/Hz}] \tag{1.24}$$

where r is the rate of the adopted channel code, and W is the reference bandwidth. The reference bandwidth can be the available bandwidth of the channel, the transmitted signal bandwidth, or any other bandwidth definition depending on the considered communication system. In a frequency division multiplexed (FDM) system, it can be defined as the distance between the carriers of two adjacent channels [17]. The achieved SE will be often placed in the Shannon plane as a function of the signal-to-noise ratio (SNR). For each SE, the corresponding SNR is the value which guarantees a reliable communication. In many practical applications the communication is considered reliable if the packet error rate (PER) is below a given threshold, for example 10^{-5}.

The second figure of merit that we consider is the *achievable information rate* (AIR). For a channel with channel law $p(\boldsymbol{r}|\boldsymbol{c})$, and a particular modulation format, the information rate (IR) is defined as [18]

$$\begin{aligned} I(\boldsymbol{c};\boldsymbol{r}) &= h(\boldsymbol{r}) - h(\boldsymbol{r}|\boldsymbol{c}) && (1.25)\\ &= \mathrm{E}\{-\log_2 p(\boldsymbol{r})\} - \mathrm{E}\{-\log_2 p(\boldsymbol{r}|\boldsymbol{c})\}\,, && (1.26)\end{aligned}$$

and measures the highest rate achievable on the channel with the adopted modulation format. In many applications, however, the receiver could consider a mismatched channel law $q(\boldsymbol{r}|\boldsymbol{c})$ (also denoted as *auxiliary channel*) different from the actual channel law $p(\boldsymbol{r}|\boldsymbol{c})$. We define thus, the *achievable information rate* as the highest rate achievable on the channel with the mismatched receiver [19, 20]. It reads

$$\begin{aligned} I_{\mathrm{R}} &= \mathfrak{h}(\boldsymbol{r}) - \mathfrak{h}(\boldsymbol{r}|\boldsymbol{c}) && (1.27)\\ &= \mathrm{E}\{-\log_2 q(\boldsymbol{r})\} - \mathrm{E}\{-\log_2 q(\boldsymbol{r}|\boldsymbol{c})\} && (1.28)\end{aligned}$$

where $q(\boldsymbol{r}) = \sum_{\boldsymbol{c}} q(\boldsymbol{r}|\boldsymbol{c})P(\boldsymbol{c})$. We point out that in (1.28) the average is computed with respect to the actual statistics $p(\boldsymbol{r}|\boldsymbol{c})$, and the mismatched entropies are explicitly denoted by $\mathfrak{h}$ to distinguish them from the standard entropies h. The AIR is always upper bounded as

$$I_{\mathrm{R}} \leq I(\boldsymbol{c};\boldsymbol{r}) \qquad (1.29)$$

with equality if and only if $q(\boldsymbol{r}|\boldsymbol{c}) = p(\boldsymbol{r}|\boldsymbol{c})$, or in other words, if and only if the receiver performs optimal detection and decoding [21].

Since the bandwidth in many applications is getting more and more a limited resource, it is of interest to evaluate the *achievable spectral efficiency* (ASE) which is defined as

$$\eta = \frac{I_{\mathrm{R}}}{TW} \quad [\mathrm{bit/s/Hz}]\,. \tag{1.30}$$

The ASE, in simple words, is the maximum SE that can be achieved with a joint detection and decoding scheme. It can be computed independently of the adopted coding scheme [21], avoiding long PER simulations.

Chapter 2

Channel shortening

THE complexity of the optimal detection increases exponentially with the memory taken into account by the detector (see §1.3). Thus, for many practical communication schemes, optimal detection can be prohibitive, since the complexity is unmanageable. In this chapter, we consider detectors with reduced complexity. The complexity reduction techniques can be classified mainly in two families:

- techniques that perform detection on the original trellis but only a fraction of the available paths is explored (e.g. the $\mathscr{M}$-BCJR in [22] and sphere decoding [23]).

- techniques that work on a reduced trellis which is then fully processed.

We consider the *channel shortening* (CS), a complexity reduction technique which belongs to the second family. Channel shortening is a technique originally proposed in 1972 by Falconer and Magee [1], and recently improved by Rusek and Prlja [2]. In this chapter we will first review the CS technique proposed by Rusek and Prlja and then, the previous works on CS in §2.2. In §2.3 we will derive an adaptive version of CS. The optimization of the transmit filter for CS detectors will be discussed in §2.4. Finally in §2.5 and §2.6 we extend the CS to other channels.

2.1 CS algorithm

Let us consider the discrete-time ISI channel with AWGN

$$r_k = \sum_{i=0}^{\nu} h_i c_{k-i} + w_k \tag{2.1}$$

where $\{c_k\}$ are the transmitted symbols belonging to a properly normalized M-ary constellation, $\{h_i\}_{i=0}^{\nu}$ are the ISI taps, ν is the channel memory, and $\{w_k\}$ are independent Gaussian random variables with variance N_0.

The observable $\{r_k\}$ can be filtered with a discrete-time filter matched to $\{h_i\}$. The resulting observable is the Ungerboeck observation model (1.5), having $g_i = \sum_k h^*_{k-i} h_k$. The optimal detection is performed by means of the BCJR algorithm. Using the matrix notations (1.8), the channel law can be expressed as

$$p(\boldsymbol{r}|\boldsymbol{c}) \propto \exp\left\{ \frac{\Re\left(\boldsymbol{c}^\dagger \boldsymbol{H}^\dagger \boldsymbol{r}\right) - \boldsymbol{c}^\dagger \boldsymbol{G} \boldsymbol{c}}{N_0} \right\} \tag{2.2}$$

where $\boldsymbol{G} = \boldsymbol{H}^\dagger \boldsymbol{H}$ and is semi-positive definite. We now consider a reduced-complexity detector, which considers a mismatched channel law

$$q(\boldsymbol{r}|\boldsymbol{c}) \propto \exp\left\{ \Re\left(\boldsymbol{c}^\dagger (\boldsymbol{H}^r)^\dagger \boldsymbol{r}\right) - \boldsymbol{c}^\dagger \boldsymbol{G}^r \boldsymbol{c} \right\}, \tag{2.3}$$

where $\boldsymbol{H}^r$ is the new front end filter, named *channel shortener*, and $\boldsymbol{G}^r$ is the ISI to be set at detector, named *target response*. The superscript r denotes that they are solely considered at *receiver*, and are different from the actual $\boldsymbol{H}$ and $\boldsymbol{G}$. For simplicity the matrix $\boldsymbol{H}^r$ and $\boldsymbol{G}^r$ in (2.3) include also the noise variance N_0. Let $L \leq \nu$ the memory taken into account by the detector. Due to this constraint on the complexity, the target response must be such that

$$(\boldsymbol{G}^r)_{ij} = 0 \quad \forall |i-j| > L. \tag{2.4}$$

The matrix $\boldsymbol{G}^r$ does not need to be semi-positive definite [24].

The achievable information rate (AIR) of the mismatched detector (see §1.4) is

$$I_\mathrm{R} = \mathfrak{h}(\boldsymbol{r}) - \mathfrak{h}(\boldsymbol{r}|\boldsymbol{c}). \tag{2.5}$$

The aim of CS is, for a given L, find the $\boldsymbol{H}^r$ and $\boldsymbol{G}^r$ which maximize the AIR. Namely we want to solve the following maximization problem

$$I_{\mathrm{OPT}} = \max_{\boldsymbol{H}^r, \boldsymbol{G}^r} I_{\mathrm{R}} \tag{2.6}$$

under the constraint (2.4).

The optimization for achievable information rate was completely solved in [2] under the assumption that $\boldsymbol{c}$ are independent Gaussian symbols. Closed-form expressions for $\boldsymbol{G}^r$, $\boldsymbol{H}^r$ for the ISI channel can be found with the following algorithm:

- Compute the sequence $\{b_i\}_{i=-L}^{L}$ as

$$b_i = \frac{1}{2\pi}\int_{-\pi}^{\pi} \frac{N_0}{|H(\omega)|^2 + N_0} e^{j\omega i} \mathrm{d}\omega \tag{2.7}$$

  where $H(\omega)$ is the discrete time Fourier transform (DTFT) of $\{h_i\}$.

- Compute the real-valued scalar

$$\mathscr{C} = b_0 - \boldsymbol{b}\boldsymbol{B}^{-1}\boldsymbol{b}^{\dagger}, \tag{2.8}$$

  where $\boldsymbol{b} = [b_1, b_2, \ldots, b_L]$, and $\boldsymbol{B}$ is $L \times L$ Toeplitz with entries $(\boldsymbol{B})_{ij} = b_{j-i}$.

- Define the vector $\boldsymbol{u} = \frac{1}{\sqrt{\mathscr{C}}}[1, -\boldsymbol{b}\boldsymbol{B}^{-1}]$ and find the optimal target response as

$$G^r(\omega) = |U(\omega)|^2 - 1, \tag{2.9}$$

  where $U(\omega)$ is the DFT of $\{u_i\}$.

- Finally, the optimal channel shortener is found as

$$H^r(\omega) = \frac{H(\omega)}{|H(\omega)|^2 + N_0}(G^r(\omega) + 1) \ . \tag{2.10}$$

By using the optimal channel shortener and the target response I_{OPT} results to be

$$I_{\mathrm{OPT}} = -\log_2(\mathscr{C}) \, . \tag{2.11}$$

The proof is shown in [2].

Clearly, when $L = \nu$, the trivial solution $G^r(\omega) = |H(\omega)|^2$, $H^r(\omega) = H(\omega)$ is found and the achievable rate simplifies to the famous formula

$$-\log_2(\mathscr{C}) = \int_{-\infty}^{\infty} \log_2\left(1 + \frac{|H(\omega)|^2}{N_0}\right) \mathrm{d}\omega \,. \tag{2.12}$$

Although the algorithm here is limited to the ISI channel, we point out that CS can be applied also to multiple-input multiple-output (MIMO) channel. In fact [2] worked on a slightly general model which represents either the ISI channel (2.1) or a the MIMO channel.

Now, the first question that the reader could ask is, why should we be interested in the optimal solution for Gaussian inputs, when practical communication schemes use finite cardinality alphabets? Although we cannot give a proof, in our experience we saw that employing the solution for Gaussian inputs for finite low-order cardinality alphabets, the resulting AIR is still excellent (see [2, 25, 26]).

Before going ahead let us show an example for the sake of clarity. We considered an EPR4 channel having channel response $\boldsymbol{h} = [0.5\,, 0.5\,, -0.5\,, -0.5]$. Figure 2.1 shows the AIR by employing a BPSK modulation and the CS detector. For comparison the figure shows also the AIR by employing the naïve technique of truncating the considered ISI at detector to L values, and the AIR for optimal detection ($L = \nu = 3$), for which the CS technique and truncation are the same algorithm. The AIR were computed by means of the Monte Carlo method described in [21]. From the figure it can be seen that CS outperforms the truncation method, with SNR gains beyond 3 dB.

2.2 Previous works on CS

The original paper in 1973 by Magee and Falconer [1] proposed channel shortening detectors optimized from a minimum mean-square-error (MMSE) perspective, and many papers in the literature followed the same approach (e.g. [27, 28]). In [1], the detector considered a mismatched channel law

$$q(\boldsymbol{r}|\boldsymbol{c}) \propto \exp\left\{-\frac{|\boldsymbol{W}\boldsymbol{r} - \boldsymbol{Q}\boldsymbol{c}|^2}{N_0}\right\} \tag{2.13}$$

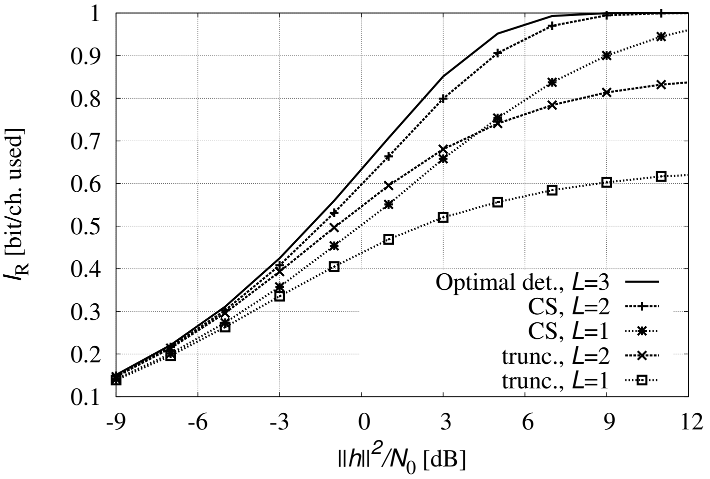

Figure 2.1: AIRs of the CS detector on the EPR4 channel.

where $\boldsymbol{W}$ and $\boldsymbol{Q}$ are Toeplitz matrix, representing the channel shortener and the target response respectively. However, two main flaws in this approach can be found: first of all, minimizing the mean-square-error does not directly correspond to achieving the highest information rate (in the Shannon sense) that can be supported by a shortening detector. Second, it can observed that (2.13) can be equivalently expressed as

$$q(\boldsymbol{r}|\boldsymbol{c}) \propto \exp\left\{ \frac{2\Re(\boldsymbol{c}^\dagger \boldsymbol{Q}^\dagger \boldsymbol{W}\boldsymbol{r}) - \boldsymbol{c}^\dagger \boldsymbol{W}^\dagger \boldsymbol{W} c}{N_0} \right\} \tag{2.14}$$

which is equal to (2.3) by setting $\boldsymbol{H}^r = \boldsymbol{W}^\dagger \boldsymbol{Q}$ and $\boldsymbol{G}^r = \boldsymbol{W}^\dagger \boldsymbol{W}$. Thus, it can be noticed that the traditional CS detector requires $\boldsymbol{G}^r$ to be semi-positive definite and $\boldsymbol{H}^r$ to have a specific structure. In [2] and [29], it is shown that the new CS presented in the previous section, outperforms the traditional CS.

Other papers on CS instead adopted other figures of merit: for example [30] proposed *maximum shortening signal-noise-ratio* (MSSNR) which minimizes the energy outside a window of interest and holds the energy inside fixed. The work in [31] instead considered the *sum-squared auto-correlation* (SA) of the combined channel-equalizer response, and tries to minimize the SA outside a window of interest. However, all these techniques do not imply a maximization of the AIR.

Before Rusek and Prlja publication [2] in 2012, in almost forty years of CS, the only work who adopted an information theoretic approach was [32] by Abou-Faycal and Lapidoth, presented in a conference in 2000.

Unfortunately, although [32] is an excellent work and has very similar results to [2], the full version of the paper was not available on the web and it had very few citations. In fact, Rusek and Prlja were not aware of [32] and worked independently. The common points and differences between the two works can be summarized as follow:

- Both works choose the channel shortener and target ISI by maximizing the achievable information rate. The target ISI $\boldsymbol{G}^r$ is found with the same algorithm (although it looks different in the two works).

- The work [32] considered the Forney detection instead of Ungerboeck detection. Thus, if $\boldsymbol{G}^r$ is positive definite, the two techniques are the same. In the case $\boldsymbol{G}^r$ is not positive definite, [32] does not give a clear answer on how the channel shortener can be computed.

- The framework of [2] was also for MIMO channel.

- The proof of [2] gave a closed formula for the channel shortener. Instead [32] left a parameter to be optimized in the formula.

2.3 Adaptive CS detector for unknown channels

The optimal CS detector in §2.1 is carried out by assuming a perfect knowledge of the ISI channel. In this section, we consider the case of unknown ISI taps $\{h_i\}$ of the channel model (2.1), and we will derive an adaptive CS detector.

The derivation of the adaptive CS detector relies on the following observations. First, the optimal channel shortener (2.10) is the combination of two filters: a MMSE filter and a filter with frequency response $G^r(\omega)+1$. Second, the sequence $\{b_i\}_{i=-L}^{L}$ is the autocorrelation of the error

$$b_i = \mathrm{E}\{e_{i+k}e_k^*\} \tag{2.15}$$

being $e_k = c_k - \hat{c}_k$ and $\hat{c}_k$ the output of the MMSE filter. This is found by observing that $N_0/(|H(\omega)|^2+N_0)$ is the power spectral density of the error at the output of a MMSE filter [33].

The adaptive CS detector can be summarized in the following steps:

- the transmitter sends a training sequence, known at receiver side.
- the receiver computes the MMSE filter by means of the training sequence.
- the receiver estimates the error correlation b_k.

The first two steps can be easily done by means of the least mean square (LMS) algorithm, or the recursive least square (RLS) algorithm [34]. The last step, the error-correlation estimation, can be easily done by computing at receiver the error sequence $\{e_k\}$, and using standard estimators (e.g., the *xcorr* in Matlab).

2.4 Optimized transmit filter for CS detector

In §2.1 we showed the algorithm to derive the optimal CS detector when the considered memory L is lower then the actual memory ν.

In this section, we extend the CS algorithm by designing a proper transmit filter to be employed jointly with a channel-shortening detector with the aim of further improving the achievable information rate. In other words, we consider to adopt, at the receiver side, a channel-shortening detector and then solve for the optimal transmit filter to be used jointly with it. When the use of the optimal full-complexity receiver is allowed, the answer to this question is the classical waterfilling filter. We are generalizing the waterfilling concept to the case of reduced-complexity channel-shortening detectors, i.e., we essentially redo Hirt's derivations [35], but this time with the practical constraint of a given receiver complexity.

Our results are not as conclusive as in the unconstrained receiver complexity case. With functional analysis, we can prove that, for real channels, the optimal transmit filter has a frequency response described by $L+1$ real-scalar values. In general, for complex channels, the optimal transmit filter is described by $L+1$ complex scalar values. The transmit filter optimization thereby becomes a problem of finite dimensionality, and a numerical optimization provides the optimal spectrum. Note that, in practice, L is limited to rather small values and $L=1$ is an appealing choice from a complexity perspective. This essentially leads to very effective numerical optimizations.

Problem formulation

We consider the channel model (2.1). The transmitted symbols $\{c_k\}$ are a precoded version of the information symbols $\{a_k\}$ as

$$\begin{aligned} c_k &= a_k \otimes p_k && (2.16) \\ &= \sum_i a_{k-i} p_i && (2.17) \end{aligned}$$

where $\{p_i\}$ is a transmit filter subject to the power constraint $\sum_i |p_i|^2 = 1$. Using a matrix notation we can express (2.16) as

$$\boldsymbol{c} = \boldsymbol{P}\boldsymbol{a} \tag{2.18}$$

where $\boldsymbol{P}$ is a Toeplitz matrix with entries $(\boldsymbol{P})_{ij} = p_{i-j}$. The combined channel-precoder thus reads

$$\begin{aligned} \boldsymbol{r} &= \boldsymbol{H}\boldsymbol{c} + \boldsymbol{w} && (2.19) \\ &= \boldsymbol{H}\boldsymbol{P}\boldsymbol{a} + \boldsymbol{w} && (2.20) \\ &= \boldsymbol{V}\boldsymbol{a} + \boldsymbol{w}, && (2.21) \end{aligned}$$

where $\boldsymbol{V} = \boldsymbol{H}\boldsymbol{P}$. Equivalently, (2.21) can be expressed by means of the scalar notation

$$r_k = \sum_{i=0}^{v_C} v_i a_{k-i} + w_k \tag{2.22}$$

where $v_i = h_i \otimes p_i$, and ν_C is the combined memory. If the CS detector with memory L is used for detection on the combined channel precoder $\boldsymbol{HP}$, the AIR (for Gaussian symbols) reads

$$I_{\mathrm{OPT}} = -\log_2(\mathscr{C}) \tag{2.23}$$

where $\mathscr{C}$ is the real-valued scalar (2.8), function of the coefficients $\{b_i\}_{i=-L}^{L}$ which read

$$\begin{aligned} b_i &= \frac{1}{2\pi}\int_{-\pi}^{\pi} \frac{N_0}{|V(\omega)|^2 + N_0} e^{j\omega i} \mathrm{d}\omega & (2.24)\\ &= \frac{1}{2\pi}\int_{-\pi}^{\pi} \frac{N_0}{|P(\omega)|^2|H(\omega)|^2 + N_0} e^{j\omega i} \mathrm{d}\omega\,. & (2.25)\end{aligned}$$

The problem we aim at solving is to maximize the AIR I_{OPT} of (2.23) over the transmit filter $P(\omega)$, i.e., the DTFT of $\boldsymbol{p}$. Thus, we have the following optimization problem at hand

$$\begin{gathered} \min_{P(\omega)} \mathscr{C}[P(\omega)] \\ \text{such that} \\ \int_{-\pi}^{\pi} |P(\omega)|^2 \mathrm{d}\omega = 2\pi \quad . \end{gathered} \tag{2.26}$$

In (2.26), we have made explicit the dependency of $\mathscr{C}$ on $P(\omega)$, but not on N_0 and $H(\omega)$, since these are not subject to optimization. Since the starting point is the expression of the AIR when the optimal channel-shortening detector is employed, we are thus jointly optimizing the channel shortening filter, the target response, and the transmit filter, although for Gaussian inputs only. However, as shown in the numerical results, when a low-cardinality discrete alphabet is employed, a significant performance improvement is still observed (see also [2]).

The optimization problem (2.26) is an instance of calculus of variations. We have not been able to solve it in closed form, but we can reduce the optimization problem into an $L+1$ dimensional problem, which can then efficiently be solved by standard numerical methods. The main result of this optimization is the following theorem.

Theorem 1. *The optimal transmit filter with continuous spectrum for the channel $H(\omega)$ with a memory L channel-shortening detector satisfies*

$$|P(\omega)|^2 = \max\left(0, \frac{N_0}{\sqrt{|H(\omega)|^2}}\sqrt{\sum_{\ell=-L}^{L} A_\ell e^{j\ell\omega}} - \frac{N_0}{|H(\omega)|^2}\right), \tag{2.27}$$

where $\{A_\ell\}$ are complex-valued scalar constants with Hermitian symmetry, i.e.

$$A_\ell = A_{-\ell}^{*}. \tag{2.28}$$

For the proof see the Appendix C.

Theorem 1 gives a general form of the optimal transmit filter to be used for a memory L channel shortening detector. By definition, it becomes the classical waterfilling filter when $L = \nu_C$. Hence, it also provides an insight to the behavior of the transmit filter for the classical waterfilling algorithm. We remind the reader that $\nu + 1$ denotes the duration of the channel impulse response and $\nu_C + 1$ denotes the duration of the combined transmit filter and channel response. We summarize our finding in the following

Theorem 2. *Let $P(\omega)$ be the transmit filter found through the waterfilling algorithm. Then,*

$$\nu_C \geq \nu.$$

For a proof, see the Appendix D.

Whereas the statement is trivial when the transmit filter and the channel have a finite impulse response (FIR), the theorem proves that this fact holds also when they have infinite impulse responses (IIR). Thus, for a FIR channel response, the waterfilling solution cannot contain any pole that cancels a zero of the channel, while, for IIR channels, the waterfilling solution cannot contain any zero that cancels a pole. Thus, the overall channel cannot be with memory shorter than the original one.

Theorem 2 reveals the interesting fact that the waterfilling algorithm trades a rate gain for detection complexity. By using the optimal transmit filter, a capacity gain is achieved, but the associated decoding complexity (of a full complexity detector) must inherently increase. Thus, with waterfilling, it is not possible to achieve both a rate gain and a decoding complexity reduction at the same time.

Numerical results for the optimized filter

Theorem 1 provides a general form of the optimal transmit filter for channel shortening detection of ISI channels. What remain to be optimized are the $L+1$ complex-valued constants $\{A_\ell\}$. A closed form optimization seems out of reach since the constraint in (2.26) has no simple analytical form in $\{A_\ell\}$. In fact, the integral

$$\int \sqrt{1+A\cos(x)}dx$$

is an instance of the incomplete elliptic integral of the second kind, for which no closed form is known to date.

We have applied a straightforward numerical optimization of the variables $\{A_\ell\}$ under the constraints in (2.26) and

$$\sum_{\ell=-L}^{L} A_\ell e^{j\ell\omega} \geq 0. \tag{2.29}$$

With a standard workstation and any randomly generated channel impulse response, the optimization is stable, converges to the same solution no matter the starting position as long as the signal-to-noise-ratio (SNR) is not very high or very low, and is altogether a matter of fractions of a second.

We now describe some illuminating examples. In all cases, the transmit power is the same both in the absence and presence of the optimal transmit filter. We first consider the complex channel $\boldsymbol{h} = [0.5, 0.5, -0.5, -0.5j]$ with memory $\nu = 3$.[1] Fig. 2.2 shows the AIR I_{OPT} for Gaussian inputs when the transmit filter is optimized for different values of the memory L considered by the receiver. For comparison, the figure also gives I_{OPT} for a flat transmit power spectrum (i.e., no transmit filter at all) and the channel capacity (i.e., when using the spectrum obtained by means of the waterfilling algorithm and assuming a receiver with unconstrained complexity). It can be seen that using an optimized transmit filter for each L, significant gains are achieved w.r.t. the flat power spectrum at all SNRs. The flat spectrum reaches its maximum information rate when $L = \nu$ but suffers a loss from the channel capacity. On the other hand,

[1] Other examples can be found in [26].

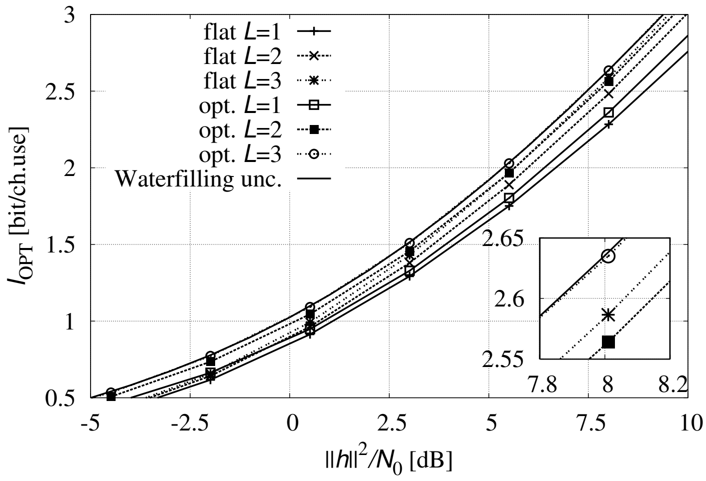

Figure 2.2: AIRs for Gaussian inputs when different values of the memory L are considered at receiver.

we can see that the optimized transmit filter when $L = \nu$ achieves an achievable rate which is close to the channel capacity. However, there is not an exact match. This loss is due to the fact that ν must be lower than the combined channel-precoder memory ν_C as stated by Theorem 2.

This behavior is clearly illustrated by Fig. 2.3, which plots the information rate when the transmit filter is found through the waterfilling algorithm and the receiver complexity is constrained with values of the memory L. It can be seen that when the memory L is increased more and more, even above ν, the information rate becomes closer and closer to the channel capacity. Moreover, it is important to notice that if, naïvely, a transmit filter found through the waterfilling algorithm is used when the receiver complexity is constrained, a loss w.r.t. the optimized case occurs and it may even be better to not have any transmit filter at all for high SNR values.

Although the results were so far presented only for Gaussian symbols, we now show that when the optimized transmit filter and detector for Gaussian inputs are used

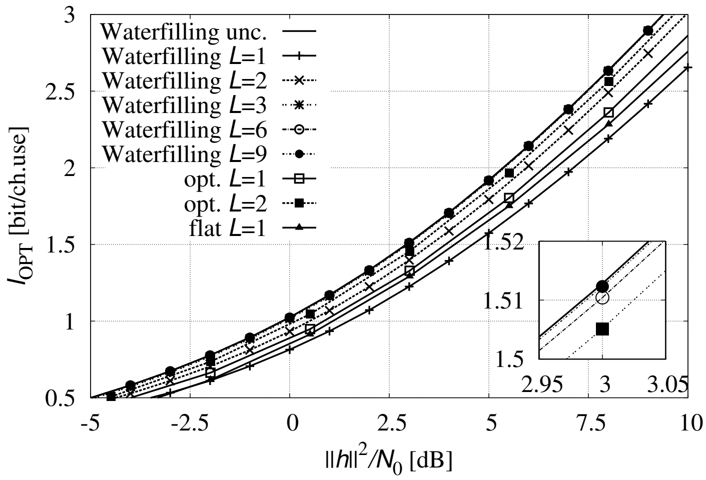

Figure 2.3: AIRs for Gaussian inputs with the waterfilling-solution power spectrum, when different values of the memory L are considered at receiver.

for low-cardinality discrete alphabets, the ensuing I_{R} is still excellent.[2] Fig. 2.4 shows the AIR for a binary phase shift keying (BPSK) modulation. It can be noticed that the behavior among the curves for BPSK reflects the behavior for Gaussian symbols. The AIR can be approached in practice with proper modulation and coding formats. Fig. 2.5 shows the bit error rate (BER) of a BPSK-based system using the DVB-S2 low-density parity-check code with rate 1/2. In all cases, 10 internal iterations within the LDPC decoder and 10 global iterations were carried out. It can be noticed that the performance is in accordance with the AIR results. All simulations that we have presented were also carried out for other channels (e.g., EPR4, Proakis B and C) and our findings for those channels are in principle identical to those for the channel here presented.

[2] We remind the reader that I_{OPT} refers to an optimized detector while I_{R} refers to the achievable rate for a non optimized detector. Since the filters have been optimized for Gaussian inputs, but we are using here low-cardinality constellations, the filters could be further optimized and for these reason we use the notation I_{R}.

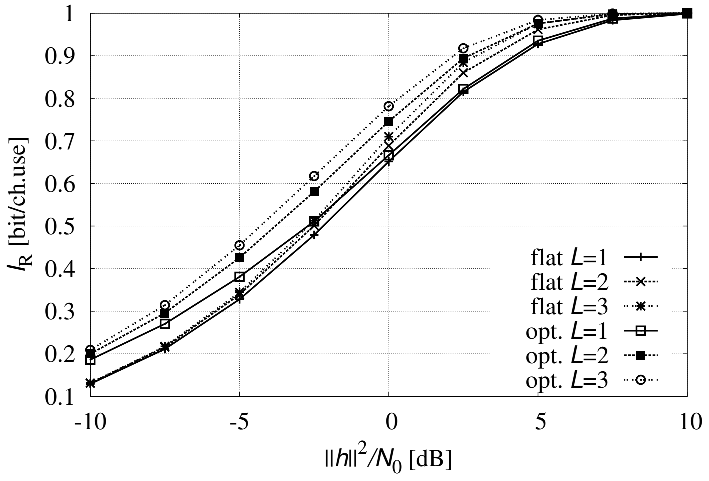

Figure 2.4: AIRs for BPSK modulation when different values of the memory L are considered at receiver.

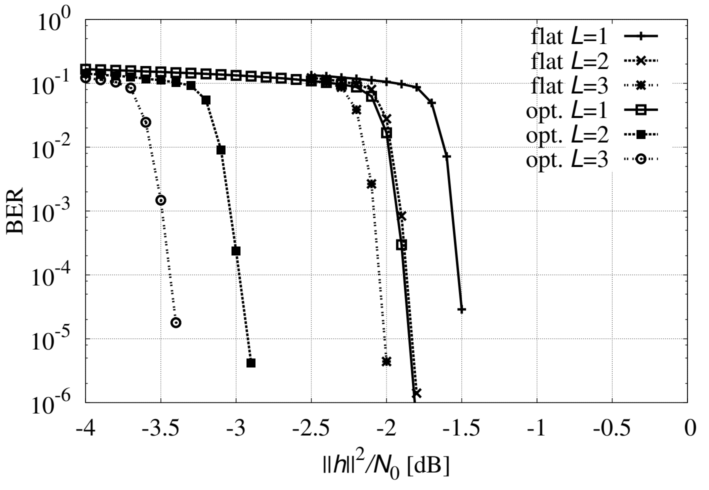

Figure 2.5: Bit error rate for BPSK modulation for different values of the memory L considered at receiver.

Finally we show that Theorem 1, similarly to waterfilling, has a graphical interpretation, although not effective as well. Let us define $A(\omega) = \sqrt{\sum_{\ell=-L}^{L} A_\ell e^{j\ell\omega}}$. It can be seen from (2.27) that $|P(\omega)|^2 \neq 0$ when $A(\omega) \geq \frac{1}{|H(\omega)|}$. Let us now consider as an example the Proakis B channel, for which $\boldsymbol{h} = [0.407, 0.815, 0.407]$ and suppose that we are constrained at the receiver side to $L = 1$. Since the channel is real, $A(\omega)$ can be expressed as a function of the two real parameters A_0 and A_1 as $A(\omega) = \sqrt{A_0 + 2A_1 cos(\omega)}$. The magnitude of $\frac{1}{|H(\omega)|}$, is depicted in Figure 2.6 (top). If we also report the optimal expression of $A(\omega)$ (the optimal coefficients are $A_0 \simeq 7.3$ and $A_1 \simeq 5.2$ when $N_0 = 0.9$) in the same figure, we obtain that the non-zero part of the frequency response of the transmit filter is given by the gray filled difference in the figure. The final amount of power spent on each frequency will be obtained by weighing the gray filled curve by $N_0/|H(\omega)|$ (center part of the figure), and by reporting it over the abscissa (bottom part of the figure).

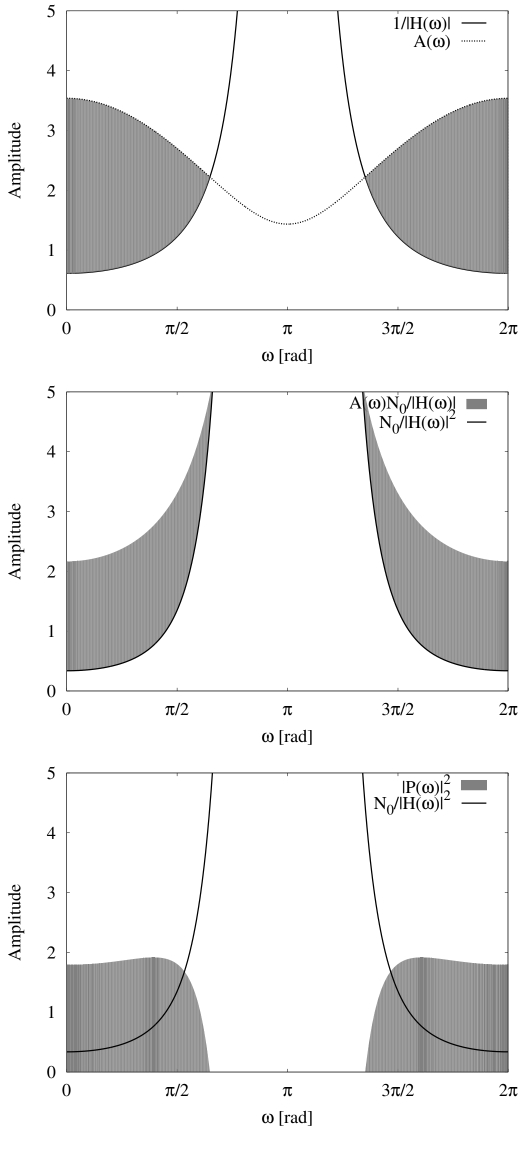

Figure 2.6: Graphical interpretation for the derivation of the optimal transmit filter for the case of the Proakis B channel, with $N_0 = 0.9$, and receiver constrained to memory $L = 1$.

2.5 Extension to MIMO-ISI channels

We extend now the CS framework to MIMO-ISI channels of the form

$$\boldsymbol{r}_k = \sum_{i=0}^{\nu} \boldsymbol{H}_i \boldsymbol{c}_{k-i} + \boldsymbol{w}_k \tag{2.30}$$

where the ISI taps $\{\boldsymbol{H}_i\}_{i=0}^{\nu}$ are $K \times K$ matrix. Vectors $\boldsymbol{r}_k$, $\boldsymbol{c}_k$, and $\boldsymbol{w}_k$ are, respectively, the observable, the transmitted symbols (independent identically distributed, IID) and the white noise having autocorrelation function

$$\mathrm{E}\left\{\boldsymbol{w}_{k+i}\boldsymbol{w}_k^{\dagger}\right\} = N_0 \boldsymbol{I} \delta_i \tag{2.31}$$

where $\boldsymbol{I}$ is the identity matrix, and δ_i the Kronecker delta. All vectors are column vectors with size K. Without loss of generality in (2.30) we assumed that the number of transmitting antennas is equal to the number of receiving antennas, since any non-square channel, can be decomposed into a square equivalent channel by means of the QR factorization [2, 29, 36].

All vectors in (2.30) can be gathered in a block-matrix notation as

$$\mathbf{r} = \mathbf{H}\mathbf{c} + \mathbf{w} \tag{2.32}$$

where $\mathbf{H}$ is block Toeplitz, with submatrix $(\mathbf{H})_{\ell m} = \boldsymbol{H}_{\ell - m}$. Notice that for $K = 1$, (2.32) becomes the matrix notation (1.10) of the scalar case. For the matrix response $\{\boldsymbol{H}_i\}$, we define the discrete time Fourier transform (DTFT) as

$$\boldsymbol{H}(\omega) = \sum_i \boldsymbol{H}_i \mathrm{e}^{-j\omega i} \tag{2.33}$$

which is equivalent to take scalar DTFT of each entry in the matrix $\{\boldsymbol{H}_i\}$. The anti-trasform is thus define as

$$\boldsymbol{H}_i = \frac{1}{2\pi}\int_{-\pi}^{\pi} \boldsymbol{H}(\omega)\mathrm{e}^{j\omega i}\mathrm{d}\omega\,. \tag{2.34}$$

The optimal detection for the channel (2.30) can be done with a generalized version of the Ungerboeck BCJR algorithm described in §1.3. The Ungerboeck observation model is derived by filtering the samples $\{\boldsymbol{r}_k\}$ with a filter matched to the

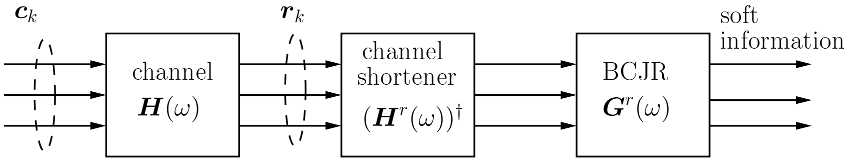

Figure 2.7: Block diagram of the CS transceiver scheme over a MIMO-ISI channel for $K = 3$.

channel, having DTFT $\boldsymbol{H}^\dagger(\omega)$. Then, BCJR detection is performed based on the metric

$$\Lambda_k(\boldsymbol{c}_k, \boldsymbol{\sigma}_k, \boldsymbol{\sigma}_{k+1}) = \exp\left\{ \frac{\Re\left(2\boldsymbol{c}_k^\dagger \boldsymbol{r}_k\right) - \boldsymbol{c}_k^\dagger \boldsymbol{G}_0 \boldsymbol{c}_k - 2\boldsymbol{c}_k^\dagger \sum_{i=1}^{\nu} \boldsymbol{G}_i \boldsymbol{c}_{k-i}}{N_0} \right\} \mathscr{I}(\boldsymbol{c}_k, \boldsymbol{\sigma}_k, \boldsymbol{\sigma}_{k+1}), \tag{2.35}$$

where $\boldsymbol{\sigma_k} = [\boldsymbol{c}_{k-1}, \ldots, \boldsymbol{c}_{k-\nu}]$ is a block vector, and $\{\boldsymbol{G}_i\}_{i=-\nu}^{\nu}$ reads

$$\boldsymbol{G}_i = \sum_{k=\max(0,i)}^{\min(\nu,\nu+i)} \boldsymbol{H}_{k-i}^\dagger \boldsymbol{H}_k . \tag{2.36}$$

Optimal detection has complexity $\mathscr{O}(M^{(\nu+1)K})$. We consider instead a channel shortening detector as in Figure 2.7 where the front-end $\boldsymbol{H}^r(\omega)$ is a matrix filter with size $K \times K$. The detection is performed on a target ISI $\{\boldsymbol{G}_i^r\}_{i=-L}^{L}$, being $L \leq \nu$ the memory taken into account at detector. The proposed CS detector performs detection on a shorter ISI, but by fully processing each matrix ISI tap $\boldsymbol{G}_i^r$ with size $K \times K$. The ensuing complexity of the CS detector is $\mathscr{O}(M^{(L+1)K})$.

The optimal front-end filter $\{\boldsymbol{H}_i^r\}$ and target response $\{\boldsymbol{G}_i^r\}$ are obtained in closed form through the following steps:

- Compute

$$\boldsymbol{B}(\omega) = N_0 \boldsymbol{H}^\dagger(\omega) \left[\boldsymbol{H}(\omega)\boldsymbol{H}^\dagger(\omega) + N_0 \boldsymbol{I}\right]^{-1} (\boldsymbol{H}^\dagger(\omega))^{-1} . \tag{2.37}$$

Applying the anti trasform to $\boldsymbol{B}(\omega)$ yields the matrix sequence $\{\boldsymbol{B}_i\}$.

- Find
$$\boldsymbol{\mathcal{C}} = \boldsymbol{B}_0 - \underline{\mathbf{B}}\mathbf{B}^{-1}\underline{\mathbf{B}}^{\dagger} \tag{2.38}$$
where we defined the block matrix $\underline{\mathbf{B}} = [\boldsymbol{B}_1, ..., \boldsymbol{B}_L]$ and the block Toeplitz $\mathbf{B}$ constructed on $\{\boldsymbol{B}_i\}$ as
$$\mathbf{B} = \begin{pmatrix} \boldsymbol{B}_0 & \boldsymbol{B}_1 & \dots & \boldsymbol{B}_{L-1} \\ \boldsymbol{B}_1^{\dagger} & \boldsymbol{B}_0 & \dots & \boldsymbol{B}_{L-2} \\ \vdots & & \ddots & \vdots \\ \boldsymbol{B}_{L-1}^{\dagger} & \boldsymbol{B}_{L-2}^{\dagger} & \dots & \boldsymbol{B}_0 \end{pmatrix} . \tag{2.39}$$

- Define the sequence $\{\boldsymbol{U}_i\}$ where $\boldsymbol{U}_0$ is the Cholesky decomposition of $\boldsymbol{\mathcal{C}}$, namely $\boldsymbol{\mathcal{C}} = \boldsymbol{U}_0^{\dagger}\boldsymbol{U}_0$, and $\boldsymbol{U}_i$ for $1 \leq i \leq L$ is the $(1,i)$ matrix entry of
$$\underline{\mathbf{U}} = -\boldsymbol{U}_0\underline{\mathbf{B}}\mathbf{B}^{-1} . \tag{2.40}$$

- Set
$$\boldsymbol{G}_i^r = \sum_{k=\max(0,i)}^{\min(L,L+i)} \boldsymbol{U}_{k-i}^{\dagger}\boldsymbol{U}_k - \delta_i \boldsymbol{I} . \tag{2.41}$$

- The optimal front-end filter is given by
$$\boldsymbol{H}^r(\omega) = \left[\boldsymbol{H}(\omega)\boldsymbol{H}^{\dagger}(\omega) + N_0\boldsymbol{I}\right]^{-1} \boldsymbol{H}(\omega)\left(\boldsymbol{G}^r(\omega) + \boldsymbol{I}\right) . \tag{2.42}$$

The proof is given in Appendix B.

Optimization of the transmit filter for MIMO-ISI channels

On MIMO-ISI channels, a transmit filter can be adopted with the aim of further improving the performance (as did for the scalar case in §2.4). Namely, we consider the transmitted symbols $\{\boldsymbol{c}_k\}$, a precoded version of the information symbols $\{\boldsymbol{a}_k\}$. We will show that the advantages of a transmit filter are twofold: we can further improve the achievable information rate, and detection can be performed as for K independent parallel channels with complexity $\mathcal{O}(M^{L+1})$.

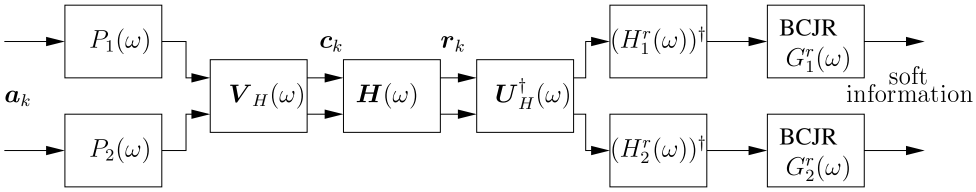

Figure 2.8: Block diagram of the transceiver for 2 × 2 MIMO-ISI channels.

It is well known that a $K \times K$ MIMO channel can be decomposed into K independent parallel channels by means of singular value decomposition (SVD) [37]. With a similar approach, the DTFT of $\{\boldsymbol{H}_i\}$ in (2.34), can be factorized by means of SVD as

$$\boldsymbol{H}(\omega) = \boldsymbol{U}_H(\omega)\boldsymbol{\Sigma}(\omega)\boldsymbol{V}_H^\dagger(\omega),$$

where $\boldsymbol{U}_H(\omega)$ and $\boldsymbol{V}_H(\omega)$ are unitary matrices and $\boldsymbol{\Sigma}(\omega)$ is a diagonal matrix with elements $\{\Sigma_i(\omega)\}_{i=1}^K$. By adopting the MIMO filter $\boldsymbol{V}_H(\omega)$ at the transmitter and the filter $\boldsymbol{U}_H^\dagger(\omega)$ at the receiver, without any information loss we obtain K independent parallel channels with channel responses $\{\Sigma_i(\omega)\}_{i=1}^K$. The transceiver block diagram is as shown in Fig. 2.8 for the case $K = 2$. The objective function to be maximized is

$$I_{\text{OPT}} = \sum_{i=1}^{K} -\log_2(\mathscr{C}_i) \tag{2.43}$$

under the constraint

$$\sum_{i=1}^{K} \frac{1}{2\pi}\int_{-\pi}^{\pi} |P_i(\omega)|^2 \mathrm{d}\omega = K \tag{2.44}$$

where $\mathscr{C}_i$ is given in (2.8) and $P_i(\omega)$ is the precoder for the channel $\Sigma_i(\omega)$. By solving the Euler-Lagrange equation, the optimal precoders have spectra of the form (2.27).

Numerical results for MIMO-ISI channels

We now consider a 2×2 MIMO-ISI channel, with $\nu = 3$ and taps

$$\boldsymbol{H}_0 = \begin{pmatrix} -0.080302 & 0.256280 \\ 0.385964 & 0.353422 \end{pmatrix} \tag{2.45}$$

$$\boldsymbol{H}_1 = \begin{pmatrix} 0.440662 & -0.168631 \\ 0.159813 & -0.338684 \end{pmatrix} \tag{2.46}$$

$$\boldsymbol{H}_2 = \begin{pmatrix} -0.358555 & -0.303972 \\ -0.084969 & 0.668917 \end{pmatrix} \tag{2.47}$$

$$\boldsymbol{H}_3 = \begin{pmatrix} 0.669006 & 0.066229 \\ 0.347376 & -0.207065 \end{pmatrix}. \tag{2.48}$$

Fig. 2.9 shows the AIR I_{OPT} for Gaussian inputs as a function of E_H/N_0, being $E_H = \sum_\ell \text{Tr}(\boldsymbol{H}_\ell \boldsymbol{H}_\ell^\dagger)$. The transmit filters are optimized for the equivalent channels $\Sigma_1(\omega)$ and $\Sigma_2(\omega)$ for different values of the memory L considered by the receiver. For comparison, the figure also gives I_{OPT} for flat transmit power spectra (i.e., $\boldsymbol{c}_k = \boldsymbol{a}_k$ and $\text{E}\{\boldsymbol{a}_{k+i}\boldsymbol{a}_k^\dagger\} = \boldsymbol{I}\delta_i$ (where $\boldsymbol{I}$ is the identity matrix and δ_i is the Kronecker delta) and the channel capacity (i.e., when using the spectra obtained by means of the waterfilling algorithm and assuming a receiver with unconstrained complexity). It can be seen that conclusions for scalar ISI channels also hold for MIMO-ISI. However, we found that, for MIMO-ISI channels, the objective function seems to have some local maxima, and thus the optimization can depend on the starting position. This problem can be easily solved by running the optimization more times (three times were always enough in all our tests) and keeping the maximum value.

2.6 Channel shortening for continuous-time channels

This section shows that the CS framework, although derived for discrete-time channel models, can be easily extended to continuous-time AWGN channels. Namely we extend the channel shortening technique to the single carrier scenario, and a multicarrier scenario, with frequency division multiplexing (FDM).

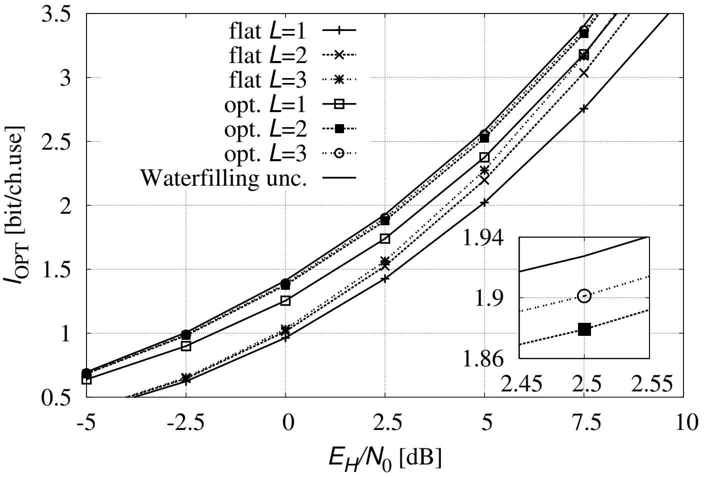

Figure 2.9: AIRs for Gaussian inputs over a MIMO-ISI channel with $K = 2$ and $\nu = 3$, when different values of the memory L are considered at the receiver.

Linear modulation on the continuous-time AWGN channel

We consider a linear modulation over a continuous-time AWGN channel. The received signal reads

$$r(t) = \sum_{k=0}^{N-1} c_k \tilde{p}(t-kT) + w(t) \tag{2.49}$$

where $\{c_k\}_{k=0}^{N-1}$ are the N transmitted symbols, which are independent and identically distributed (IID). The $\tilde{p}(t)$ is the shaping pulse, T the symbol time, and $w(t)$ is white Gaussian noise with power spectral density N_0. The shaping pulse $\tilde{p}(t)$ is constrained to have bandwidth W and energy

$$\int_{-W/2}^{W/2} |\tilde{P}(f)|^2 \mathrm{d}f = 1 \tag{2.50}$$

being $\tilde{P}(f)$ the Fourier transform of $\tilde{p}(t)$. The channel is assumed perfectly known at the receiver and time-invariant. The channel frequency response is assumed flat over W, although the generalization to the case of a frequency-selective channel is straightforward.

As explained in §1.2, a sufficient statistics for detection of (2.49) can be carried out by using a whitening matched filter (WMF). The ensuing observable, is the Forney observation model and reads as (2.22), where the ISI $\{v_i\}$ of the combined channel precoder is such that

$$|V(\omega)|^2 = \frac{1}{T}\sum_i \left|\tilde{P}\left(\frac{\omega}{2\pi T} - \frac{i}{T}\right)\right|^2 , \tag{2.51}$$

and optimal detection can be performed with the BCJR algorithm (see §1.3). Clearly, this discrete-time model will depend on the adopted shaping pulse, its bandwidth, the employed symbol time, and the channel impulse response if the channel is frequency selective.

Instead of optimal detection, we want to consider a CS detector with memory L. The optimal target ISI and channel shortener are derived again through (2.8)–(2.10) by means of the $\{b_i\}_{i=-L}^{L}$ in (2.24). The corresponding channel shortening receiver is shown in Figure 2.10a. Since the WMF can be implemented as a cascade of a continuous-time matched filter followed by a discrete-time whitening filter, this latter filter can be combined with the channel shortening filter obtaining a single discrete-time filter with frequency response

$$\tilde{H}^r(\omega) = \frac{G^r(\omega)+1}{|V(\omega)|^2 + N_0} . \tag{2.52}$$

The corresponding channel shortening receiver is shown in Figure 2.10b.

The shaping pulse $\tilde{p}(t)$ can be also optimized by means of the framework in §2.4. The DTFT of $\{v_k\}$ can be decomposed as

$$|V(\omega)|^2 = |P(\omega)|^2 |H(\omega)|^2 \tag{2.53}$$

where

$$H(\omega) = \begin{cases} 1 & |\omega| \leq 2WT\pi \\ 0 & \text{otherwise} \end{cases} , \ \omega \in [-\pi, \pi] . \tag{2.54}$$

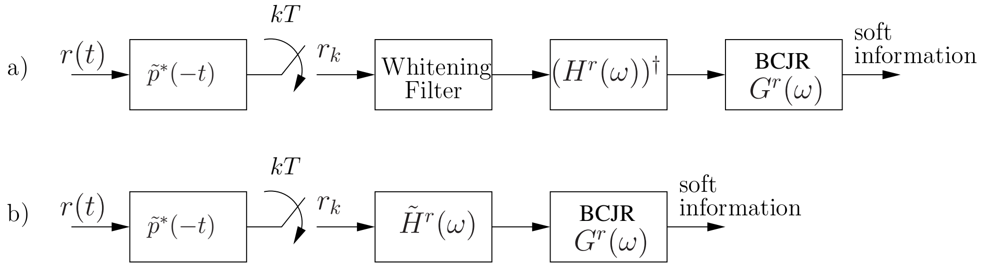

Figure 2.10: Block diagram of the CS receiver for continuous-time AWGN channels.

Thus the optimization problem is still given by (2.26) where the optimal shaping pulse is such that

$$|\tilde{P}(f)|^2 = T|P(2\pi Tf)|^2 \tag{2.55}$$

with $|P(\omega)|^2$ given in (2.27).

Clearly, when $2WT \geq 1$, the optimal solution is trivial and $|P(\omega)|^2$ is flat. Thus, for $2WT = 1$ the $\tilde{p}(t)$ is a sinc function, whereas for $2WT > 1$ the $\tilde{p}(t)$ can be a pulse whose spectrum has vestigial symmetry (e.g., pulses with a root raised cosine (RRC) spectrum). For $2WT < 1$, the symbol time is such that the Nyquist condition for the absence of ISI cannot be satisfied. Thus, we are working in the domain of the *faster-than-Nyquist* (FTN) paradigm [3, 38, 39] or its extension represented by time packing [5, 40]. Note that, as said before, the discrete-time channel model, will depend on the values of W and T. When changing the values of W and/or T, the corresponding optimal pulse will change and so the maximum value of the AIR for the given allowed complexity. In general, when reducing the value of WT, the maximum AIR value will decrease. However, the spectral efficiency, defined as the ratio between the AIR and the product WT could, in principle, increase [3, 38, 41, 39, 40, 5]. This is the rationale behind FTN/time packing that allows to improve the spectral efficiency by accepting interference. The optimal value of T is, in that case, properly optimized to maximize the spectral efficiency. This optimization can be now performed by also using, for each value of T, the corresponding optimal shaping pulse. In other words, we can find the optimal pulse for a constrained complexity detector when FTN/time packing is adopted.

We point out that, for this scenario, the numerical computation of the optimal shaping pulse in the time-domain can require the adoption of some windowing technique or the use of Parks-McClellan algorithm [42] to obtain a practical pulse since $H(\omega)$ has a spectrum with an ideal frequency cut.

We finally point out that the optimization of the shaping pulse to frequency selective AWGN channels is done straightforwardly by properly defining the channel (2.54).

Numerical results for the optimized shaping pulse

We computed the optimal shaping pulse on a bandlimited AWGN channel when the bandwidth W and the symbol time T are such that $2WT = 0.48$. Hence, we are in the realm of FTN/time packing and the considered ISI is only due to the adoption of such a technique. Fig. 2.11 shows the achievable spectral efficiency (ASE) $\eta = I_{\mathrm{R}}/WT$ for a BPSK modulation on the continuous-time AWGN channel as a function of the ratio E_b/N_0, E_b being the received signal energy per information bit. Two values of the memory, namely $L = 1$ and $L = 2$ are considered at the detector. For comparison, the figure also gives the ASE for pulses with RRC spectrum and roll-off $\alpha = 0.1$ or $\alpha = 0.2$, and the unconstrained capacity for the AWGN channel. It can be seen that the optimized pulse outperforms the other pulses.

FDM on the continous-time AWGN channel

We consider a scenario with K carriers, each transmitting the symbols $\{c_k^{(\ell)}\}_{k=0}^{N-1}$, being ℓ the index of the carrier. The received signal reads

$$r(t) = \sum_{\ell=0}^{K-1} \sum_{k=0}^{N-1} c_k^{(\ell)} p_\ell(t-kT) e^{j2\pi F_\ell t} + w(t) \tag{2.56}$$

where F_ℓ is the frequency of the ℓ-th carrier, $p_\ell(t)$ its shaping pulse, T the symbol time, and $w(t)$ white Gaussian noise with power spectral density N_0.

A sufficient statistics $\boldsymbol{r}_k = [r_k^{(0)}, \ldots, r_k^{(K-1)}]^T$, for the detection of (2.56) is found by adopting a bank of matched filter to the received signal. The block diagram for

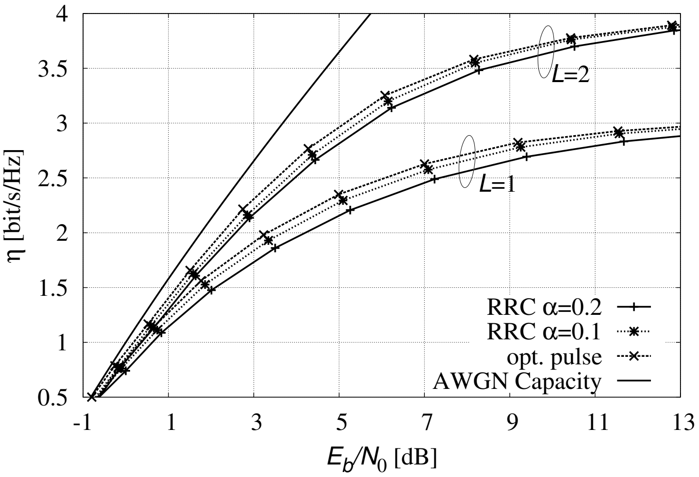

Figure 2.11: ASE for a BPSK modulation by using the optimized pulse for two values of the memory L considered at receiver.

$K = 3$ is shown in Figure 2.12. It can be shown that the observable $\boldsymbol{r}_k$ reads

$$\boldsymbol{r}_k = \sum_i \boldsymbol{G}_{k,k-i}\boldsymbol{c}_{k-i} + \boldsymbol{n}_k \tag{2.57}$$

where $\boldsymbol{c}_k = [c_k^{(0)}, \dots, c_k^{(K-1)}]^T$, $\boldsymbol{n}_k$ is colored Gaussian noise with autocorrelation

$$\mathrm{E}\{\boldsymbol{n}_{k+i}\boldsymbol{n}_k\} = N_0\boldsymbol{G}_{k+i,k} \tag{2.58}$$

and $\boldsymbol{G}_{k,j}$ is a $K \times K$ matrix with entries

$$\left(\boldsymbol{G}_{k,j}\right)_{\ell,u} = e^{-j2\pi(F_\ell - F_u)jT} \int_{-\infty}^{\infty} p_u(t)p_\ell^*(t-(k-j)T)e^{-j2\pi(F_\ell - F_u)t}\mathrm{d}t\,. \tag{2.59}$$

Clearly the channel (2.57) is not stationary, since the matrix $\boldsymbol{G}_{k,j}$ are time variant. However if the receiver is modified as in Figure 2.13 we obtain the equivalent stationary channel

$$\boldsymbol{z}_k = \sum_i \tilde{\boldsymbol{G}}_i\boldsymbol{x}_{k-i} + \tilde{\boldsymbol{n}}_k \tag{2.60}$$

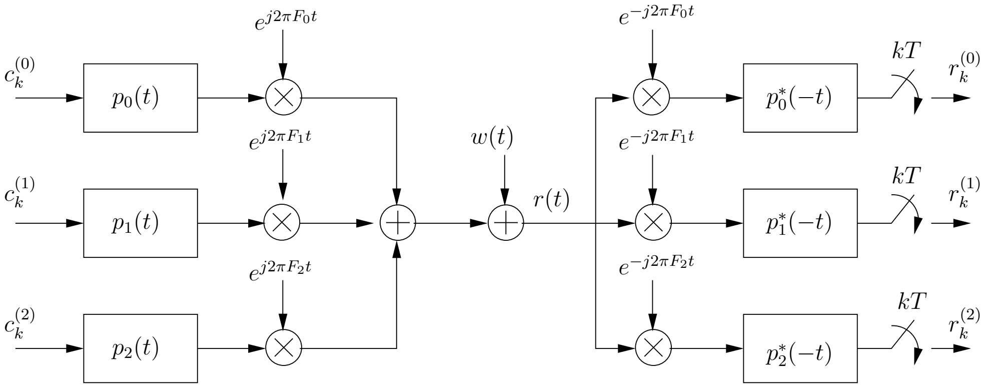

Figure 2.12: Block diagram of the FDM transceiver scheme over the AWGN channel, when $K = 3$.

where

$$\left(\tilde{\boldsymbol{G}}_i\right)_{\ell,u} = e^{-j2\pi(F_\ell - F_u)iT} \int_{-\infty}^{\infty} p_u(t) p_\ell^*(t - iT) e^{-j2\pi(F_\ell - F_u)t} \mathrm{d}t\,, \tag{2.61}$$

$$\boldsymbol{x}_k = \boldsymbol{c}_k \circ \begin{bmatrix} e^{j2\pi F_0 kT} \\ \vdots \\ e^{j2\pi F_{K-1} kT} \end{bmatrix}, \tag{2.62}$$

and $\circ$ is the Hadamard product. The equivalent stationary channel does not involve any information loss. In fact, it is easy to prove that

$$I(\boldsymbol{c};\boldsymbol{r}) = I(\boldsymbol{c};\boldsymbol{z})\,. \tag{2.63}$$

Finally, it can be noticed that (2.60) is the Ungerboeck observation model of a MIMO-ISI channel. Thus the CS detector can be designed as described in §2.5.

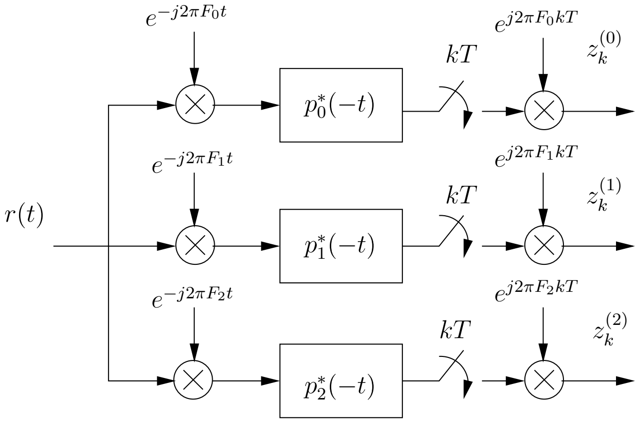

Figure 2.13: Modified receiver to obtain an equivalent stationary FDM channel.

Chapter 3

Time packing

IN satellite links for broadcasting and broadband applications, orthogonal signaling, that ensures absence of intersymbol interference (ISI), is often adopted. As an example, in the 2nd-generation satellite digital video broadcasting (DVB-S2) standard [43], a conventional square-root raised-cosine (RRC) pulse shaping filter is specified at the transmitter. In an additive white Gaussian noise channel and in the absence of other impairments, the use of a matched filter (MF) at the receiver and proper sampling ensure that optimal detection can be performed on a symbol-by-symbol basis. On the other hand, it is known that, when finite-order constellations are considered [e.g., phase-shift keying (PSK)], the efficiency of the communication system can be improved by giving up the orthogonality condition, thus accepting interference. For example, *faster-than-Nyquist signaling* (FTN, see [3, 4]) is a well known technique consisting of reducing the spacing between two adjacent pulses in the time-domain well below the Nyquist rate, thus introducing ISI. If the receiver is able to cope with the interference, the efficiency of the communication system will be increased. In the original papers on FTN signaling [3, 4], this optimal time spacing is obtained as the smallest value giving no reduction of the minimum Euclidean distance with respect to the Nyquist case. This ensures that, asymptotically, the ISI-free performance is reached, at least when the optimal detector is adopted. The i.u.d. capacity or information rate, i.e., the average mutual information when the channel inputs are independent and uniformly distributed (i.u.d.) random variables, is then computed, still assuming the adoption of the optimal detector [44, 45]. However, the complexity of this optimal detector easily becomes unmanageable, and no hints are provided on how to perform the optimization in the more practical scenario where a reduced-complexity receiver is employed.

In [5], a different approach for improving the spectral efficiency, that relies on both *time packing* of adjacent symbols and reducing the spacing of the adjacent channels when applicable (multi-carrier transmission), has been considered. It is assumed that, at the receiver side, a symbol-by-symbol detector working on the samples at the MF output is adopted, and the corresponding information rate is computed, by also optimizing time and frequency spacings to maximize the *achievable spectral efficiency* (ASE). Hence, rather than the minimum distance, the *ASE is the performance measure* and, in addition, a low-complexity detection algorithm, characterized by a given allowable complexity *irrespectively of the interference set size*, is considered at the receiver rather than the optimal detector employed in [3, 4, 44, 45]. Although the MF output represents a set of sufficient statistics for optimal detection, a suboptimal symbol-by-symbol receiver is considered in [5]. Hence, the ASE can be improved by employing more sophisticated detection algorithms. In this chapter, we will consider two cases: (i) a proper filtering of the MF output plus a symbol-by-symbol detector and (ii) the maximum *a posteriori* (MAP) symbol detector that, in order to limit the receiver complexity, takes into account only a limited amount of interference.

This technique represents a good alternative, for low-order constellations, to the shaping of the transmitted symbol distribution [46], providing spectral efficiencies that cannot be reached when orthogonal signaling is employed. Improving the ASE without increasing the constellation order can be considerably convenient since the larger the constellation size, the higher the decoding complexity and the lower the robustness to channel impairments such as time-varying phase noise and non-linearities. In the case of frequency packing, a further improvement could be achieved by adopting, at the receiver side, a multi-user detector. The remainder of this chapter is organized as follows. The system model is described in §3.1. In §3.2, we compute and optimize the spectral efficiency considering detectors with different complexity. Numerical results are reported in §3.4, where we also show the performance of some efficient modulation and coding formats (MODCODs) designed accordingly.

3.1 System model

We consider an additive white Gaussian noise (AWGN) channel and a frequency-division multiplexed system where perfectly synchronized (downlink assumption) adjacent channels employ the same linear modulation format, shaping pulse $p(t)$, and symbol interval (or time spacing) T. The shaping pulse is assumed to have unit energy. The received signal can be expressed as

$$r(t) = \sqrt{E_s}\sum_{k,\ell} c_k^{(\ell)} p(t-kT)e^{j2\pi\ell Ft} + w(t) \tag{3.1}$$

where E_s is the symbol energy, $c_k^{(\ell)}$ the symbol transmitted over the u-th channel during the ℓ-th symbol interval, F the frequency spacing between adjacent channels, and $w(t)$ a circularly symmetric zero-mean white Gaussian noise process with power spectral density N_0. The transmitted symbols $\{c_k^{(\ell)}\}$ are independent and uniformly distributed and belong to a given zero-mean M-ary complex constellation χ properly normalized such that $E\{|c_k^{(\ell)}|^2\} = 1$. Note that the summations in (3.1) extend from $-\infty$ to $+\infty$, namely an infinite number of time epochs and carriers are employed. For the spectral efficiency computation, we will consider the central user only using F as a measure of the signal bandwidth.

The base pulse $p(t)$ has often RRC-shaped spectrum (RRC pulse in the following) with roll-off factor α. In addition to it, we will consider other transmit pulses, e.g., a pulse whose spectrum is raised-cosine (RC) shaped (RC pulse in the following) and a Gaussian pulse. In general, we will consider the case of time-frequency packing and we will optimize the frequency separation F between two adjacent users and the symbol interval T in order to maximize the ASE. In the case of bandlimited pulses (i.e., RRC and RC pulses), we will also consider time packing only. In this case, adjacent users are not allowed to overlap in frequency (i.e., $F = (1+\alpha)/T$ for RRC and RC pulses) and we may assume that only the user with $\ell = 0$ is transmitted. In satellite communications, this can correspond to the use of a single carrier occupying the entire transponder bandwidth. This is of particular interest since the on-board power amplifier can operate closer to saturation and hence improve the efficiency.

3.2 Spectral efficiency optimization

In this section, we shown how to compute the ASE for a given receiver and how to optimize the values of T and F.

Symbol-by-Symbol detection

Let us consider the central user (i.e., that for $\ell = 0$). We first consider the case shown in Figure 3.1(a) of a receiver composed by a filter matched to the shaping pulse $p(t)$, followed by a proper discrete-time filter, that works on $\gamma \geq 1$ samples per symbol interval, and a symbol-by-symbol (SBS) detector. Although the discrete-time filter could be, in general, fractionally-spaced (FS, i.e., $\gamma > 1$), the detector will operate on one sample per symbol interval. These samples will be denoted by $\{r_k^{(0)}\}$ and can be expressed as

$$r_k^{(0)} = \sqrt{E_s} c_k^{(0)} h(0,0,k) + \sqrt{E_s} \sum_{(n,\ell) \neq (0,0)} c_{k-n}^{(\ell)} h(n,\ell,k) + z_k \qquad (3.2)$$

in which $h(n,\ell,k)$ is the residual interference at time kT due to the ℓ-th user and the $(k-n)$-th transmitted symbol, and $\{z_k\}$ is the additive noise term, in general colored unless a whitening filter (WF) is employed after the MF. The discrete-time filter is assumed properly normalized such that the noise variance is N_0. The dependence of coefficients $h(n,\ell,k)$ on k is through a complex coefficient of unit amplitude which disappears for $\ell = 0$ (hence $h(n,0,k)$ is independent of k) and is due to the fact that F is not an integer multiple of $1/T$.

Eq. (3.2) shows the two different impairments experienced by the receiver, namely the background noise and the interference. Instead of simply neglecting the interference due to adjacent symbols and users, we pursue here a more general approach, which consists of modeling the interference as a zero-mean Gaussian process with

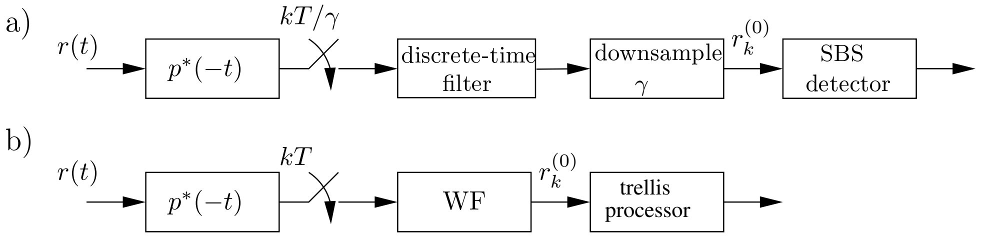

Figure 3.1: Some considered receivers: (a) symbol-by-symbol detector and (b) single-user detector based on trellis processing.

power spectral density equal to N_I, of course independent of the additive thermal noise—we point out that this approximation is exploited only by the receiver, while in the actual channel the interference is clearly generated as in (3.2). Note that the interference is really Gaussian distributed only if the transmitted symbols $c_k^{(\ell)}$ are Gaussian distributed as well. However, especially when the interference set is small, e.g., when T and F are large, the actual interference distribution may substantially differ from a Gaussian distribution.

We define *auxiliary channel* the channel model assumed by the receiver. With the above mentioned Gaussian approximation, the auxiliary channel is

$$r_k^{(0)} = \sqrt{E_s} c_k^{(0)} h(0,0,k) + \mathfrak{v}_k \tag{3.3}$$

where $\{\mathfrak{v}_k\}$ are independent and identically distributed zero-mean circularly symmetric Gaussian random variables, with variance $N_0 + N_I$. It turns out that

$$N_I = E_s \sum_{(n,\ell)\neq(0,0)} |h(n,\ell,k)|^2 \tag{3.4}$$

which results to be independent of k, as can be easily shown. We are interested in evaluating the ultimate performance limits achievable by a symbol-by-symbol receiver designed for the auxiliary channel (3.3) when the actual channel is that in (3.2), in terms of information rate (or spectral efficiency). This issue is an instance of *mismatched* detection [19] (see also [21]). The achievable information rate (AIR), measured in bit per channel use, for this mismatched receiver is

$$I_{\mathrm{R}} = \mathfrak{h}(r_k^{(0)}) - \mathfrak{h}(r_k^{(0)}|c_k^{(0)}) \tag{3.5}$$

where

$$\begin{aligned} \mathfrak{h}(r_k^{(0)}) &= -\mathrm{E}\left\{\log_2\left(\sum_{c\in\chi} q(r_k^{(0)}|c)\frac{1}{M}\right)\right\} && (3.6)\\ \mathfrak{h}(r_k^{(0)}|c_k^{(0)}) &= -\mathrm{E}\left\{\log_2 q(r_k^{(0)}|c_k^{(0)})\right\} && (3.7)\end{aligned}$$

where $q(r_k^{(0)}|c_k^{(0)})$ is a Gaussian probability density function (PDF) of mean $c_k^{(0)}$ and variance (N_0+N_I) (in accordance with the auxiliary channel model), while the outer statistical average, with respect to $c_k^{(0)}$ and $r_k^{(0)}$, is carried out according to the real channel model (3.2) [21]. Eq. (3.5) can be evaluated efficiently by means of a Monte Carlo average [21]. From a system viewpoint, the spectral efficiency, that is the amount of information transmitted per second and per Hertz, is a more significant quality figure than the information rate. Under the assumption of infinite transmission, the ASE is defined as

$$\eta = \frac{I_{\mathrm{R}}}{FT} \quad [\mathrm{b/s/Hz}]\,. \tag{3.8}$$

For a given constellation and shaping pulse, it is possible to find the spacings T and F that provide the largest ASE. In general, we could expect that the optimal spacings depend on the signal-to-noise ratio (SNR). In fact, it is possible to show that, as the SNR increases, not only does the ASE increase, but also the optimal values of the spacings change. The properties of the function $\eta(T,F,E_S/N_0)$ cannot be easily studied in closed form, but it is clear, by physical arguments, that it is bounded, continuous in T and F, and tends to zero when $T,F \to 0$ or $T,F \to \infty$. Hence, the function $\eta(T,F)$ has a maximum value—according to our findings, in most cases there are no local maxima other than the global maximum. Formally, for a given modulation format, shaping pulse, and value of E_S/N_0, the optimization problem consists of finding the maximum of $\eta(T,F,E_S/N_0)$ varying T and F. This problem can be solved by evaluating $\eta(T,F,E_S/N_0)$ on a grid of values of T and F (coarse search), followed by an interpolation of the obtained values (fine search).

A measure of the SNR more significant than E_s/N_0 is given by E_b/N_0, being E_b the mean energy per information bit, for which $E_s = I(E_s)E_b$ holds. The optimization problem becomes

$$\eta_{\mathrm{M}}(E_b/N_0) = \max_{T,F>0} \eta(T,F,E_b/N_0)\,. \tag{3.9}$$

In order to solve it for a given value of E_b/N_0, we employed the following technique. The AIR is first evaluated for some values of the couple (T,F), and E_s/N_0. The two sets, including their cardinalities, must be designed so as to ensure an accurate sampling of the AIR, when the latter is interpreted as a function of T, F, and E_s/N_0. For each couple (T_i,F_j), cubic spline interpolation can be used to obtain a continuous function of E_s/N_0 (fine search), denoted as $I(T_i,F_j,E_s/N_0)$. Then, given a value of E_b/N_0 the following fixed-point problems are solved in E_s/N_0 for different couples (T_i,F_j),

$$\frac{E_s}{N_0} = I\left(T_i,F_j,\frac{E_s}{N_0}\right)\frac{E_b}{N_0} \tag{3.10}$$

and the AIRs corresponding to the solutions are denoted by $I(T_i,F_j,E_b/N_0)$. Further improvements could be achieved by adding N_I as variable in eq. (3.9). However, we have found by numerical results that choosing N_I as in (3.4) is almost optimal.

The spectral efficiency depends on the employed discrete-time filter. Since the optimization of this filter with the aim of maximizing the spectral efficiency is a hard task, we restricted our analysis to the cases of a WF, that will be also considered in §3.2, and of a minimum mean square error (MMSE) feedforward equalizer, possibly fractionally spaced (FS) with at most 22 taps.

Single-User Trellis Processing

Improved, still achievable, lower bounds can be obtained by relaxing the constraint on the adopted detection algorithm. In other words, we can consider a more complex receiver able to cope with (a portion of) the interference introduced by the adoption of the time-frequency packing. The receiver considered in this section will not cope with the interference due to the adjacent users—a single-user receiver is still adopted.

For a general channel with finite intersymbol interference, an optimal MAP symbol detector can be designed working on the samples at the WF output as shown in Figure 3.1(b). These samples, denoted to as Forney observation model (see §1.2), can still be expressed as in (3.2) with a proper expression of coefficients $h(n,\ell,k)$. We assume to adopt the optimal receiver for the following auxiliary channel:

$$r_k^{(0)} = \sqrt{E_s} \sum_{0 \le n \le L} f_n c_{k-n}^{(0)} + \mathfrak{v}_k \tag{3.11}$$

where $\{f_n\}_{n \ge 0}$ are such that $f_n = h(n,0,k)$ and, as mentioned, are independent of k, whereas the noise samples $\{\mathfrak{v}_n\}$, that take into account the white noise and the residual interference, are assumed independent and identically distributed zero-mean circularly symmetric Gaussian random variables with variance $(N_0 + N_I)$, with

$$N_I = \sum_{n>L} E_s |f_n|^2 + \sum_n \sum_{\ell \neq 0} E_s |h(n,\ell,k)|^2 \,. \tag{3.12}$$

which is still independent of k. The corresponding MAP symbol detector takes the form of the classical algorithm by Bahl, Cocke, Jelinek and Raviv (BCJR) [12] working on a trellis whose state takes into account L interfering symbols only, according to a given maximal allowable receiver complexity. The number of trellis states is equal to $S = M^L$.

Let us define $\boldsymbol{c} = [c_0^{(0)}, c_1^{(0)}, ..., c_{N-1}^{(0)}]$ and $\boldsymbol{r} = [r_0^{(0)}, r_1^{(0)}, ..., r_{N-1}^{(0)}]$, N being a proper integer. The simulation-based method described in [21] allows to evaluate the AIR for the mismatched receiver, i.e.,

$$\begin{aligned} I_{\mathrm{R}} &= \lim_{N \to +\infty} \frac{1}{N} I(\boldsymbol{c};\boldsymbol{r}) \\ &= \lim_{N \to +\infty} \frac{1}{N} E\left\{ \log_2 \frac{q(\boldsymbol{r}|\boldsymbol{c})}{q(\boldsymbol{r})} \right\} \left[\frac{\text{bit}}{\text{ch. use}} \right] . \end{aligned} \tag{3.13}$$

In (3.13), $q(\boldsymbol{r}|\boldsymbol{c})$ and $q(\boldsymbol{r})$ are PDF according to the auxiliary channel model, while the outer statistical average is with respect to the input and output sequences evaluated according to the actual channel model [21]. Eq. (3.13) can be evaluated recursively through the forward recursion of the BCJR detection algorithm matched to the auxiliary channel model [21]. Once the AIR has been computed, the spectral efficiency can be derived and the optimal time and frequency spacings optimized accordingly, as described in the previous section. For channels with finite ISI, optimal MAP symbol detection can be equivalently implemented by working directly on the MF output [16], i.e., on the so-called Ungerboeck observation model (see §1.2). The equivalence does not hold when reduced-complexity detection is considered and interference from adjacent channels arises. Since it is difficult to predict which is the most convenient observation model, it is of interest to evaluate the ASE when both models are employed and this can be done as described for the Forney model (see also [47] for details).

Multi-User Detection

Although the assumption of a single-user auxiliary channel gives very useful results, tighter lower bounds can be obtained by using a more general auxiliary channel model. In fact, we can consider a receiver for the central user (that with $\ell = 0$) that, in addition to the interference taken into account by the receivers in §3.2, also takes into account the J adjacent signals on each side as well (multi-user receiver)—we again point out that this approximation is exploited only by the receiver, while in the actual channel the interference is generated as in (3.1). The exact MAP receiver for the multi-user auxiliary channel can be easily derived and employed to find the ASE in the new scenario. The benefit of employing the multi-user auxiliary channel model when evaluating the ASE is two-fold: first, it allows to evaluate the performance degradation due to the use of single-user receivers, despite the presence of a strong adjacent channel interference, with respect to a more involved multi-user receiver, which is more *matched* to the real channel. Second, it gives a practical performance upper bound when low-complexity approximate multi-user receivers, for example based on linear equalization or interference cancellation, are employed (as examples, those in [48] and references therein). Obviously, in this case some (limited) degradation must be expected.

3.3 Channel shortening detection for time packing

In the previous section, we adopted mismatched detectors which consider just a limited amount of interference. The interference considered at the detector is a truncation of the actual interference (both in time and frequency). Clearly, the truncation does not maximize the ASE, which is instead what time packing aims to. A better approach than truncation, is the adoption of channel shortening (CS) detectors, as described in §2.6. Namely, for a given memory L considered at detector, we set the ISI at detector and the front-end filter as the ones which maximize the achievable spectral efficiency. Moreover, as shown in §2.6, the shaping pulse can be also optimized to improve the performance.

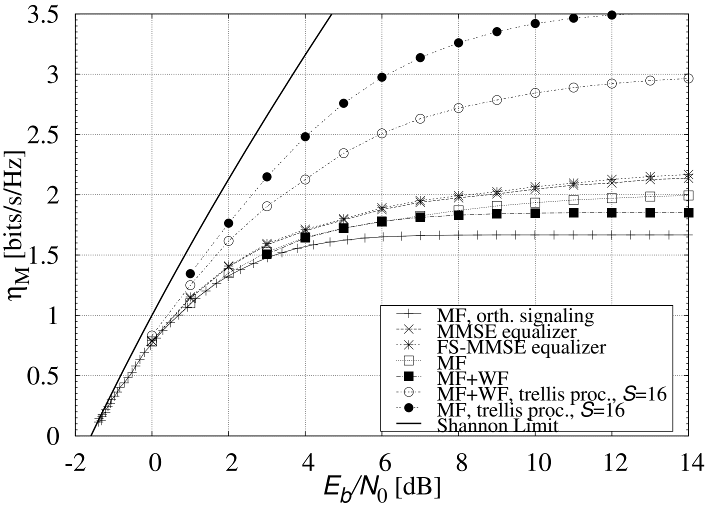

Figure 3.2: ASE for QPSK with Gray mapping and a RRC pulse having $\alpha = 0.2$.

3.4 Numerical results

In this section, we report the optimized spectral efficiency η_{M} as a function of E_b/N_0 for different modulation formats and shaping pulses. The considered modulation formats are the quaternary and octal PSK (QPSK and 8PSK).

Fig. 3.2 shows the optimized ASE in case of time packing only for the QPSK modulation with a RRC pulse with roll-off $\alpha = 0.2$. Both symbol-by-symbol detection and trellis processing (this latter taking into account $L = 4$ interfering symbols) are considered assuming Gray mapping. In this case, at the receiver side we may use two identical and independent detectors, one working on the in-phase and the other one on the quadrature component. This is beneficial in case of adoption of a MAP symbol detector. In fact, when L interfering symbols are taken into account, we have two detectors working on a trellis with 2^L states instead of a single detector working on a trellis with 4^L states. Hence, for a given complexity, a larger number of interferers can be taken into account. The curve related to the absence of time packing (i.e., in

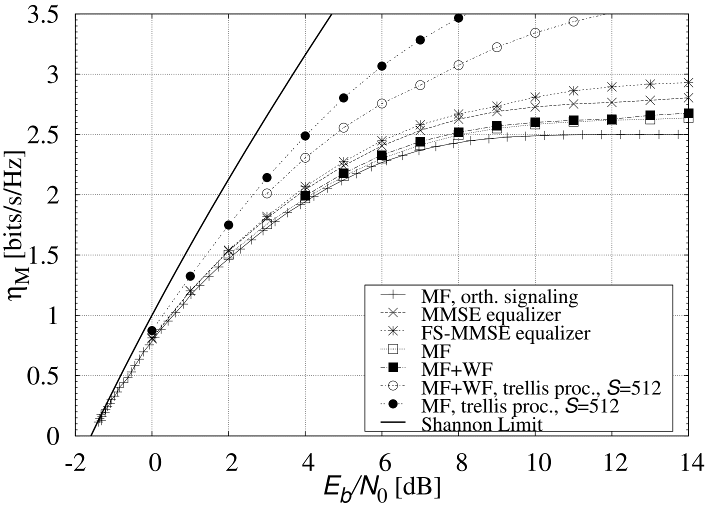

Figure 3.3: ASE for 8PSK with a RRC pulse having $\alpha = 0.2$.

case of orthogonal signaling) and the Shannon Limit for AWGN [18], are also shown for comparison. It can be observed that the time-packing technique allows to improve the spectral efficiency for each E_b/N_0 value with respect to the case of orthogonal signaling. Moreover it can noticed that, in case of use of a symbol-by-symbol detector, the FS-MMSE equalizer seems the best option whereas the Ungerboeck observation model is more suited in case of trellis processing. Similar considerations hold for the 8PSK modulation with a RRC pulse of $\alpha = 0.2$. The relevant results are shown in Fig. 3.3. Still considering QPSK with Gray mapping and trellis processing with $S = 16$, we evaluated the effect of different shaping pulses. In particular, RRC and RC pulses with different roll-off factors have been considered along with prolate spheroidal wave functions [49] and the Gaussian pulse. In these two latter cases, frequency packing is also employed. Fig. 3.4 shows the performance of some of the considered pulses. In particular, RRC pulses with α equal to 0.2 and 1.0 outperform all other pulses at low and high E_b/N_0 values, respectively. In particular, an impres-

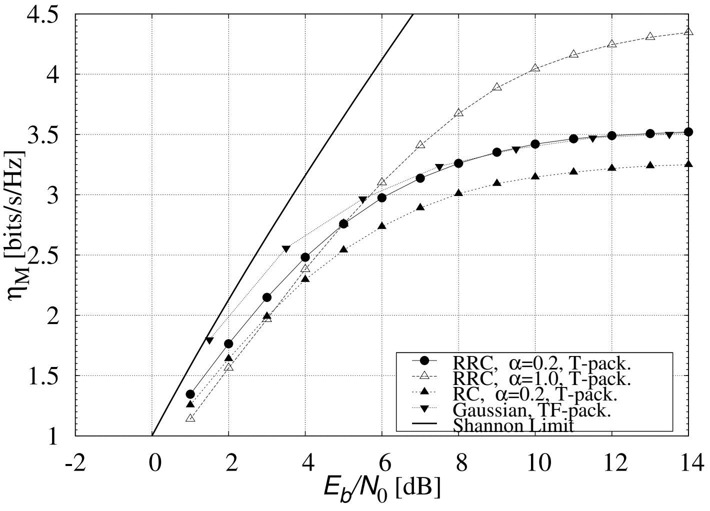

Figure 3.4: ASE for QPSK with Gray mapping by using different pulses. At the receiver, a MF front end and trellis processing with $S = 16$ is considered.

sive asymptotic spectral efficiency of 4.3 bit/s/Hz is obtained with QPSK and $\alpha = 1$.[1]

What information theory promises can be approached by using proper coding schemes. We considered MODCODs using the low-density parity-check (LDPC) codes with length 64,800 bits of the DVB-S2 standard [43], properly combined with QPSK and 8PSK modulations with time packing. RRC pulses with $\alpha = 0.2$ or $\alpha = 1$ are considered. The corresponding packet error rate (PER) have been computed by means of Monte Carlo simulations and the results are reported in the spectral efficiency plane in Fig. 3.5 using, as reference, an MPEG PER of 10^{-4}. In the same figure, the performance of the MODCODs based on the same LDPC codes with orthogonal signaling and employing QPSK, 8PSK, and the amplitude phase-shift keying (APSK) modulation with 16 and 32 symbols (16- and 32APSK) [43], are also

[1] This is due to the fact that the shaping pulse is smoother and so, for a given value of T, the introduced interference is lower.

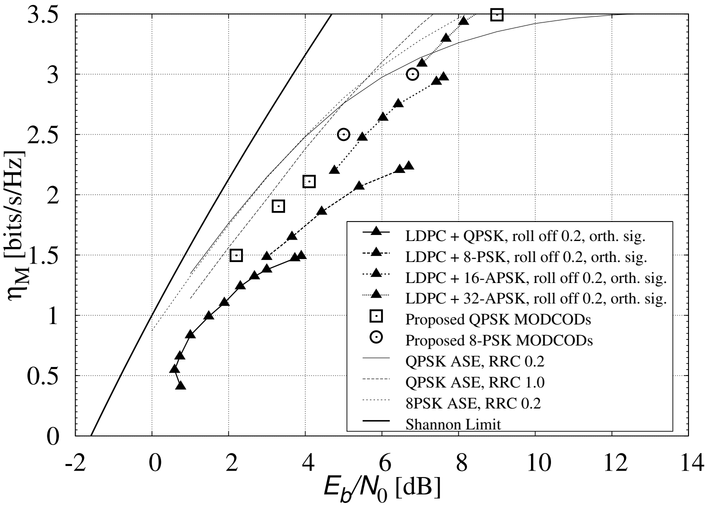

Figure 3.5: Designed MODCODs for QPSK and 8PSK with RRC pulses.

shown for comparison. We can observe that we can reach, with QPSK, values of spectral efficiency that, in case of orthogonal signaling, cannot be reached even with 16APSK.

The performance can be improved by adopting a CS detector. Figure 3.6 shows the optimized ASE for RRC and QPSK modulation with Gray mapping and $\alpha = 0.2$ by adopting the time packing technique and single-user trellis processing. Detection is performed by adopting two identical and independent detectors, one working on the in-phase and the other one on the quadrature component. The figure shows the ASE obtained by single-user trellis processing in either case of CS or truncation of the ISI. The considered numbers of states are $S = 8$ and $S = 64$. In the case of ISI truncation, we optimized also the noise variance at detector, in order to achieve the best performance. For comparison purpose, we also showed the ASE by quadrature amplitude (QAM) with cardinality 64 and 256 with orthogonal signaling. It can be noticed that CS outperforms the truncation and exhibits an excellent ASE, which for some SNR values is even higher than the one achieved by QAM modulations.

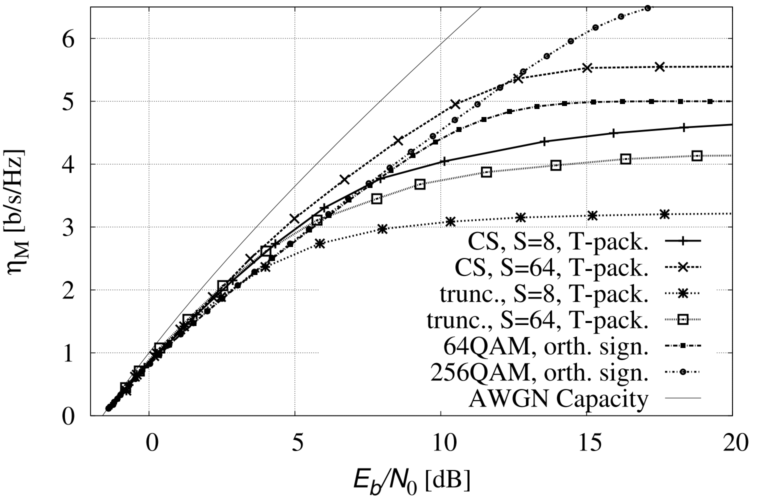

Figure 3.6: ASE for time packing and CS detection when the modulation is QPSK, with Gray mapping and RRC pulse $\alpha = 0.2$.

Chapter 4

Detection for satellite channels

SATELLITE channels are affected by nonlinear distortions and by intersymbol interference (ISI). The former originate from the presence of a high power amplifier (HPA), whereas the latter is introduced by the input and output multiplexing (IMUX and OMUX) filters placed before and after the HPA. During the last decades, the nonlinear effects and the channel memory have been coped with nonlinear compensation and data predistortion at the transmitter side (see [50] and references therein) or with advanced detection techniques (see [51] and references therein).

When the channel memory is too large to be taken into account at the detector, these advanced detection techniques quickly become unmanageable and low-complexity solutions are required. The conceptually simplest solution is to let the detector work with a truncated version of the channel response. However, as expected, such a strategy often yields poor performance unless the truncated part of the channel response has negligible power. Channel shortening, already described in Chapter 2, can be an alternative. This chapter generalizes the analysis in [2] to maximum-a-posteriori (MAP) detection for nonlinear satellite channels. In §4.1, we briefly review the system model for the satellite channel and the underlying detection algorithm assumed in this thesis. In §4.2, we extend the channel shortening technique and in §4.3 we assess its performance by numerical simulations.

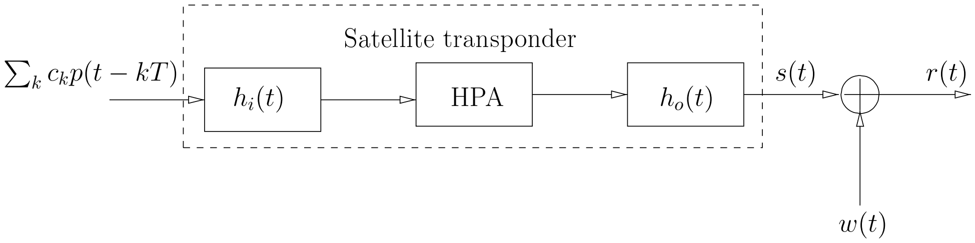

Figure 4.1: Block diagram of the satellite channel.

4.1 System model and considered detector

We consider a linear modulation with shaping pulse $p(t)$, symbol time T, and uniformly and identically distributed input symbols $\{c_k\}$ belonging to an M-ary constellation, properly normalized such that $\mathrm{E}\{|c_k|^2\} = 1$. The nonlinear satellite channel, considering a single-channel-per-transponder scenario, is depicted in Figure 4.1. It includes an IMUX filter $h_i(t)$ which removes the adjacent channels, a HPA, and an OMUX filter $h_o(t)$ aimed at reducing the spectral broadening caused by the nonlinear amplifier. Although the HPA is a nonlinear memoryless device, the overall system has memory due to the presence of IMUX and OMUX filters. The received signal is further corrupted by additive white Gaussian noise whose low-pass equivalent $w(t)$ has power spectral density N_0. The complex baseband representation of the received signal has thus the following expression

$$r(t) = s(t) + w(t)\,, \tag{4.1}$$

where $s(t)$ is the signal at the output of the OMUX filter.

In [51], it is shown that a suitable approximate model for the signal $s(t)$ is based on the following vth-order (with v being any odd integer) *simplified* Volterra-series expansion

$$s(t) \simeq \sum_k \sum_{i=0}^{N_V-1} c_k \left[|c_k|^{2i} h^{(2i+1)}(t-kT) \right], \tag{4.2}$$

where $N_V = (v+1)/2$, and $h^{(2i+1)}(t)$ are complex waveforms given by linear combinations of the the original N_V Volterra kernels. This simplified Volterra-series expansion is obtained from the classical one by neglecting some suitable terms. For further details, the reader can refer to [51]. We point out that the approximation (4.2) is used only for the receiver design and not for generating the received signal $r(t)$.

It is easy to show that MAP symbol detection based on this simplified model can be performed through a bank of filters followed by a conventional BCJR detector [12] with proper branch metrics and working on a trellis whose number of states exponentially depends on the channel memory. When the actual channel memory is large, we have to resort to complexity reduction techniques. A possible approach is the use of reduced-state techniques (e.g., see [15]) or the use of the graph-based technique described in [51] whose complexity linearly depends on the channel memory. However, to obtain a further complexity reduction, all these techniques can be combined with the CS technique described in [2] properly extended to the channel at hand.

We will separately consider the cases of phase-shift keying (PSK) modulations and amplitude/phase shift keying (APSK) modulations typically employed in satellite transmissions.

PSK modulations

It can be seen that the condition $|c_k|^2 = 1$ implies that the signal (4.2) simplifies to a linear modulation

$$s(t) \simeq \sum_k c_k \bar{h}(t - kT) \tag{4.3}$$

where $\bar{h}(t) = \sum_{i=0}^{N_V - 1} h^{(2i+1)}(t)$ [51]. In this case, detection can be perfomed using the samples $\{r_k\}$ at the output of a filter matched to $\bar{h}(t)$ as described in §1.3, and the application of CS can be carried out as described in §2.6 for linear channels.

APSK modulations

The samples at the output of a bank of filters matched to the pulses $h^{(2i+1)}(t)$, for $i = 0, \ldots, N_V - 1$ form a set of sufficient statistics for detection. Namely, considering an ν-th order expansion, we have N_V matched filters whose output, sampled at discrete time k can be collected in a $N_V \times 1$ vector that can be expressed as

$$\boldsymbol{r}_k = \sum_i \boldsymbol{G}_i \boldsymbol{c}_{k-i} + \boldsymbol{n}_k \,, \tag{4.4}$$

where $\boldsymbol{c}_k = \left[c_k,\ c_k|c_k|^2, ..., c_k|c_k|^{v-1}\right]^T$,

$$\boldsymbol{G}_i = \begin{pmatrix} g_i^{(1,1)} & g_i^{(1,3)} & \cdots & g_i^{(1,v)} \\ g_{-i}^{(1,3)*} & g_i^{(3,3)} & \cdots & g_i^{(3,v)} \\ \vdots & & \ddots & \vdots \\ g_{-i}^{(1,v)*} & g_{-i}^{(3,v)*} & \cdots & g_i^{(v,v)} \end{pmatrix} , \tag{4.5}$$

having defined $g_i^{(m,l)} = \int_{-\infty}^{\infty} h^{(n)}(t) h^{(m)*}(t - lT) \mathrm{d}t$, and $\boldsymbol{n}_k$ is a Gaussian vector with

$$\mathrm{E}\{\boldsymbol{n}_{k+i} \boldsymbol{n}_k^\dagger\} = N_0 \boldsymbol{G}_i \,. \tag{4.6}$$

Vectors $\{\boldsymbol{r}_k\}$ can be collected into a single vector

$$\mathbf{r} = \mathbf{G}\mathbf{c} + \mathbf{n} \,, \tag{4.7}$$

where $\mathbf{G}$ is a block Toeplitz matrix constructed from the matrices $\{\boldsymbol{G}_i\}$, whereas $\mathbf{c}$ and $\mathbf{n}$ are block vectors from $\{\boldsymbol{c}_k\}$ and $\{\mathbf{n}_k\}$. The channel is fully described through its conditional probability density function of the output given the input symbols, which reads

$$p(\mathbf{r}|\mathbf{c}) \propto \exp\left(\frac{2\Re(\mathbf{c}^\dagger \mathbf{r}) - \mathbf{c}^\dagger \mathbf{G} \mathbf{c}}{N_0}\right) . \tag{4.8}$$

According to the CS approach, a low-complexity detector works on a mismatched channel law [2]

$$q(\mathbf{r}|\mathbf{c}) \propto \exp\left(2\Re(\mathbf{c}^\dagger (\mathbf{H}^r)^\dagger \mathbf{r}) - \mathbf{c}^\dagger \mathbf{G}^r \mathbf{c}\right) , \tag{4.9}$$

where $\mathbf{H}^r$,$\mathbf{G}^r$ are block Toeplitz matrices constructed from the sequences $\{\boldsymbol{H}_i^r\}$ and $\{\boldsymbol{G}_i^r\}$, respectively, being $\{\boldsymbol{H}_i^r\}$ the channel shortener operating on $\mathbf{r}$, and $\{\boldsymbol{G}_i^r\}$ the target response, to be properly designed. Without loss of generality we absorb the noise variance N_0 into the two matrices in (4.9). In order to reduce the detection complexity, we constrain $\{\boldsymbol{G}_i^r\}$ to

$$\boldsymbol{G}_i^r = \mathbf{0} \quad |i| > L \tag{4.10}$$

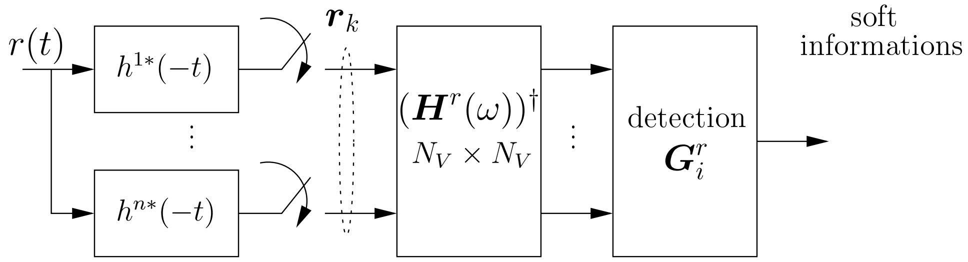

Figure 4.2: Block diagram of the suboptimal receiver for the nonlinear satellite channel.

which implies that the memory after CS is L instead of the true memory of the channel. The resulting receiver is suboptimal since it assumes (4.9) rather than the actual law (4.8), and is depicted in Figure 4.2.

4.2 Channel shortening

The achievable information rate (AIR) I_R of a mismatched detector that works with (4.9) is given by

$$\begin{aligned} I_{\mathrm{R}} &= \mathfrak{h}(\mathbf{r}) - \mathfrak{h}(\mathbf{r}|\mathbf{c}) && (4.11) \\ &= \lim_{N\to\infty} \frac{1}{N} \mathrm{E}\left\{ \log_2 \frac{q(\mathbf{r}|\mathbf{c})}{q(\mathbf{r})} \right\} \text{ [bit/ch.use]} && (4.12) \end{aligned}$$

where N is the number of transmitted symbols and the average is carried out w.r.t. **r** and **c**, according to the actual channel (see [21]).

The CS technique finds the optimal $\mathbf{H}^r$,$\mathbf{G}^r$ solving the following optimization problem

$$\arg \max_{\mathbf{H}^r,\mathbf{G}^r} I_{\mathrm{R}} \qquad (4.13)$$

under the constraints specified in (4.10).

Problem (4.13) for a discrete alphabet is a complicated task. However it can be solved in closed form under the assumption that **c** is composed of Gaussian random variables. Although this assumption is not even approximately true, since the actual symbols are functions of each other, we will show in the simulation results that a very good performance is still achieved.

Defining the correlation matrix $\boldsymbol{V} = \mathrm{E}\{\boldsymbol{c}_k\boldsymbol{c}_k^\dagger\}$, the optimal matrix-valued front-end filter $\{\boldsymbol{H}_i^r\}$ and target response $\{\boldsymbol{G}_i^r\}$ are obtained in closed form through the following steps:

- Compute the DTFT matrix $\boldsymbol{G}(\omega)$ of $\boldsymbol{G}_i$ and use the spectral decomposition to find $\boldsymbol{L}(\omega)$, i.e., decompose $\boldsymbol{G}(\omega) = \boldsymbol{L}^\dagger(\omega)\boldsymbol{L}(\omega)$. Compute

  $$\begin{aligned}\boldsymbol{B}(\omega) = {} & N_0\boldsymbol{V}\boldsymbol{L}^\dagger(\omega)\\ & \cdot\left[\boldsymbol{L}(\omega)\boldsymbol{V}\boldsymbol{L}^\dagger(\omega)+N_0\boldsymbol{I}\right]^{-1}\left(\boldsymbol{L}^\dagger(\omega)\right)^{-1}\end{aligned} \tag{4.14}$$

  where $\boldsymbol{I}$ is the identity matrix. The anti trasform yields the matrix sequence $\{\boldsymbol{B}_i\}$ having size $N_V \times N_V$.

- Find

  $$\boldsymbol{\mathcal{C}} = \boldsymbol{B}_0 - \underline{\mathbf{B}}\mathbf{B}^{-1}\underline{\mathbf{B}}^\dagger \tag{4.15}$$

  where we defined the block matrix $\underline{\mathbf{B}} = [\boldsymbol{B}_1, \ldots, \boldsymbol{B}_L]$ with size[1] $N_V \times N_V L$ and the block Toeplitz $\mathbf{B}$ with size $N_V L \times N_V L$ constructed on $\{\boldsymbol{B}_i\}$.

- Define the sequence $\{\boldsymbol{U}_k\}$ where $\boldsymbol{U}_0$ is the Cholesky decomposition of $\boldsymbol{\mathcal{C}}$, namely $\boldsymbol{\mathcal{C}} = \boldsymbol{U}_0^\dagger\boldsymbol{U}_0$, and $\boldsymbol{U}_i$ for $1 \le i \le L$ is the $(1,i)$ matrix entry of $\underline{\mathbf{U}} = -\boldsymbol{U}_0\underline{\mathbf{B}}\mathbf{B}^{-1}$.

- Set

  $$\boldsymbol{G}_i^r = \sum_{k=\max(0,i)}^{\min(L,L+i)} \boldsymbol{U}_{k-i}^\dagger\boldsymbol{U}_k - \boldsymbol{V}\delta_i \tag{4.16}$$

  where δ_i is the Kronecker delta.

- The optimal front-end filter is given by

  $$\begin{aligned}\left(\boldsymbol{H}^r(\omega)\right)^\dagger = {} & \left(\boldsymbol{G}^r(\omega)+\boldsymbol{V}^{-1}\right)\\ & \cdot\boldsymbol{V}\boldsymbol{L}^\dagger(\omega)\left[\boldsymbol{L}(\omega)\boldsymbol{V}\boldsymbol{L}^\dagger(\omega)+N_0\boldsymbol{I}\right]^{-1}\left(\boldsymbol{L}^\dagger(\omega)\right)^{-1}.\end{aligned} \tag{4.17}$$

[1] Here, the size for block matrix means the number of scalar entries, as well as for standard matrix. This notation is different from the one adopted in [25].

The algorithm is carried out by observing that the channel equation (4.7), under the Gaussian assumption, is the Ungerboeck observation model of the MIMO-ISI channel (2.32). Thus the proof is the one showed in Appendix B for MIMO-ISI channels.

We point out that by analogy to [2] for linear channels, when $L = 0$ the optimal channel shortener is a special case of the MMSE filter of [52] applied to (4.2).

4.3 Numerical results

We consider 8PSK and 16APSK modulations. The shaping pulse $p(t)$ has a root-raised-cosine (RRC) spectrum with roll-off 0.05. The IMUX and OMUX filters have frequency characteristics specified in [43] with a 3dB bandwidth of $0.94/T$ and $0.85/T$ respectively. The nonlinear transfer characteristic is the Saleh model [53] with parameters $\alpha_a = 2.1322$, $\alpha_\phi = 1.7054$, $\beta_a = 1.0746$, and $\beta_\phi = 1.5072$. A 5th-order Volterra expansion is considered at the receiver. We report all results as functions of the ratio between the normalized power at the saturation P_{sat} and the noise power spectral density N_0.

The AIR in eq. (4.12) can be computed using the Monte Carlo method described in [21]. Figure 4.3 shows the AIR values when CS is employed in combination with a 8PSK modulation, and an input back-off (IBO) equal to zero. Results are shown for different values of the detector memory L and an optimization of the noise variance at the receiver has been carried out to further improve the approximate model. For comparison, we show also the AIR values when a simple truncation of the ISI at the detector is adopted. The detector with $L = 4$ can be considered as effective as a full complexity one, since most of the ISI is taken into account[2]. It can be seen that CS has higher AIR than a simple truncation of the ISI response, even though it is designed for a vector $\mathbf{c}$ with Gaussian components. CS with memory $L = 2$ gives only a minimal performance degradation for all P_{sat}/N_0 values. Similar conclusions hold for the 16APSK, depicted in Figure 4.4. We found similar CS gains also with other modulations (QPSK and 32APSK) and other transponder characteristics (e.g.,

[2] The pulse with RRC spectrum gives an infinite memory of the channel. However, based on investigations beyond those presented in this thesis, we may assume that $L = 4$ is almost optimal.

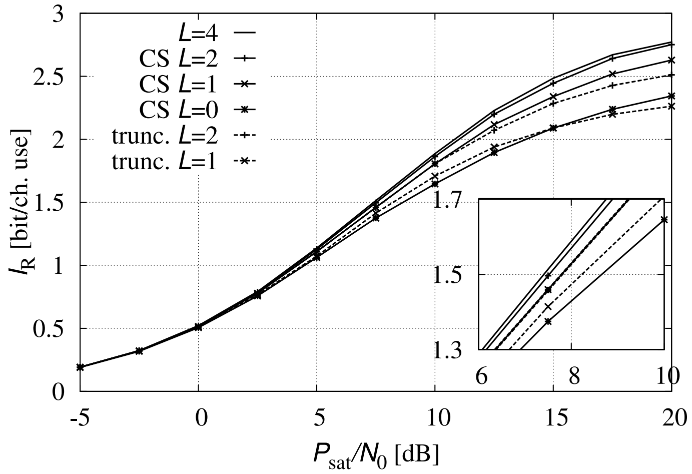

Figure 4.3: AIR for 8PSK modulation on the nonlinear satellite channel with IBO=0 dB.

the HPA in [43]).

The AIRs can be approached in practice with proper modulation and coding (MODCODs) formats. In Figure 4.5 we report the bit error rate (BER) of some MODCODs based on the DVB-S2 low-density parity-check code (LDPC) with rate 1/2. We performed iterative detection and decoding with a maximum of 50 global iterations. We note that the MODCODs performance reflects the AIRs well.

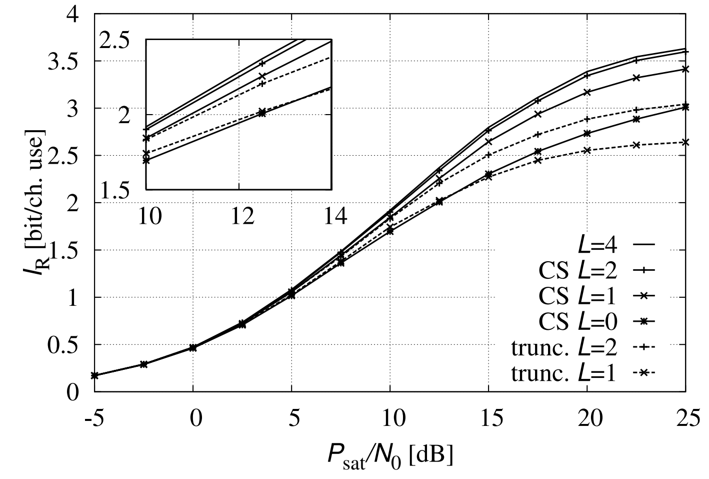

Figure 4.4: AIR for 16APSK modulation on the nonlinear satellite channel with IBO=3 dB.

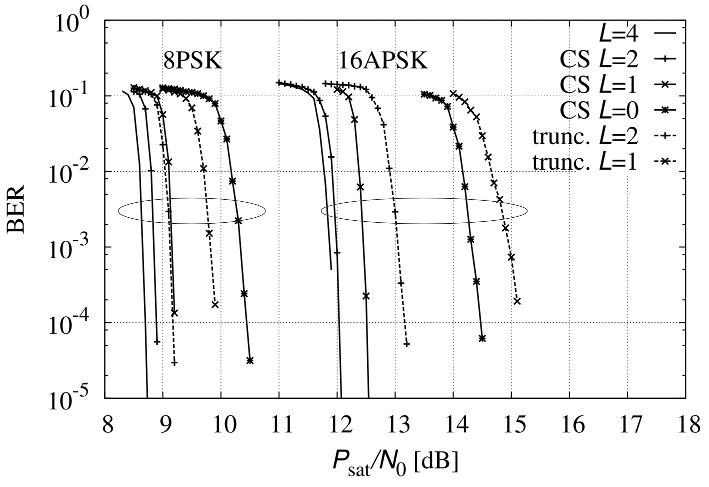

Figure 4.5: BER for 8PSK and 16APSK modulation, with the DVB-S2 LDPC code having rate 1/2.

4.4 Conclusions

We generalized the CS technique to the case of nonlinear satellite channels. We showed that when the memory L is lower than the channel memory, an optimization of the mismatched channel law at the detector yields significantly better performance than a truncation of the channel impulse response.

Chapter 5

Spectrally efficient communications over the satellite channel

In this chapter, we apply the time-frequency packing (TF packing) technique to nonlinear satellite channels. In particular, we design highly efficient schemes by choosing the time and frequency spacings which give the maximum value of SE. We assume a realistic satellite channel where nonlinear distortions originate from the presence of a high-power amplifier (HPA). The considered system is also affected by intersymbol interference (ISI), due to the presence of input and output multiplexing (IMUX and OMUX) filters placed before and after the HPA and intentionally introduced by the adoption of the time packing technique as well. We limit our investigation to systems in single-carrier-per-transponder operation (i.e., each transponder is devoted to the amplification of the signal coming from only one user[1]). Although interchannel interference (ICI) is also present due to frequency packing of signals coming from different transponders, a single-user detector is considered at the receiver. We consider two different approaches to detection for nonlinear channels, namely the use of a detector taking into account the nonlinear effects and a more traditional

[1] However, the actual transponders can be hosted on two or more co-located satellites.

scheme based on predistortion and memoryless detection. In the case of predistortion, we consider the dynamic data predistortion technique described in [54, 55], whereas in the case of advanced detection we consider the receiver described in the previous chapter employing the channel shortening technique (CS) for nonlinear satellite channels. It should be noted that we apply the TF packing technique to nonlinear satellite channels for which, usually, even by using RRC pulses there is still ISI at the receiver.

As mentioned, with respect to Chapter 3, we here consider nonlinear satellite channels instead of linear ones. Our aim here is to show the benefits that can be obtained by employing time-frequency packing jointly with a channel shortening receiver.

The proposed TF packing technique promises to provide increased SEs at least for low-order modulation formats. In fact, when dense constellations with shaping are employed, we fall in a scenario similar to that of the Gaussian channel with Gaussian inputs for which orthogonal signaling with no excess bandwidth (rectangular shaping pulses) is optimal (although this is mainly true for the linear channel and not in the presence of a nonlinear HPA, since shaping increases the peak-to-average power ratio). Improving the achievable SE without increasing the constellation order can be considerably convenient since it is well known that low-order constellations are more robust to channel impairments such as time-varying phase noise and non-linearities. It is expected that the use of low-order modulations in conjunction with TF packing provides similar advantages in terms of robustness against channel impairments.

The proposed approach to improve the SE is very general and the case of satellite systems for broadband and broadcasting applications must be thus considered just as an example to illustrate the benefits that can be obtained through the application of the TF packing paradigm coupled with advanced receiver processing.

The remainder of this chapter is organized as follows. In §5.1, we introduce the system model. The framework that we use to evaluate the SE of satellite systems is detailed in §5.2, whereas different approaches to the detection for the considered channel are described in §5.3. Numerical results are reported in §5.4, where we show how the proposed technique can improve the SE of DVB-S2 systems. Finally, conclusions are drawn in §5.5.

5.1 System Model

We consider the forward link of a transparent satellite system, where synchronous users employ the same linear modulation format, shaping pulse $p(t)$, and symbol interval (or time spacing) T, and access the channel according to a frequency division multiplexing scheme. The transmitted signal in the uplink can be expressed as

$$x(t) = \sum_{\ell} \sum_{k} c_k^{(\ell)} p(t - kT) e^{j2\pi \ell F_u t}\,, \tag{5.1}$$

where $c_k^{(\ell)}$ is the symbol transmitted by user ℓ during the k-th symbol interval, and F_u is the frequency spacing between adjacent channels.[2] The transmitted symbols belong to a given zero-mean M-ary complex constellation. Notice that, in order to leave out border effects, the summations in (5.1) extend from $-\infty$ to $+\infty$, namely an infinite number of time epochs and carriers are considered. In DVB-S2 standard the base pulse $p(t)$ is an RRC-shaped pulse with roll-off factor α (equal to 0.2, 0.3, or 0.35 depending on the service requirements). We denote by W the bandwidth of pulse $p(t)$. In case of pulses employed in DVB-S2, it is $W = (1+\alpha)/T$ since orthogonal signaling is considered. When time packing is adopted, W becomes a further degree of freedom, as described later.

As commonly assumed for broadband and broadcasting systems, on the feeder uplink (between the gateway and the satellite) the impact of thermal noise can be neglected due to a high transmit signal strength. Hence, in our analysis, we have considered a noiseless feeder uplink. Although the TF packing can be applied to other and more general scenarios, we consider here a single-carrier-per-transponder scenario, where different carriers undergo independent amplification by different transponders on board of satellite, each of which works with a single carrier occupying its entire bandwidth. This case is particularly relevant for digital broadcasting services since it allows a more efficient use of on-board resources (in particular the HPA can work

[2] In this scenario, we will use the terms, “channels”, “users”, and “carriers” interchangeably.

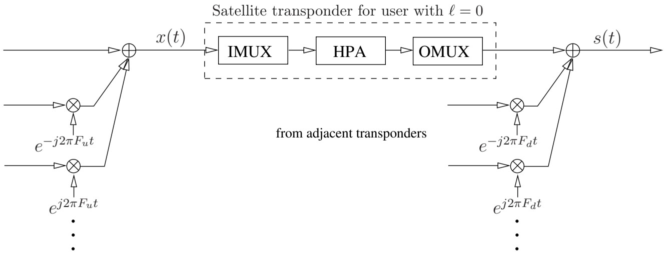

Figure 5.1: System model.

closer to saturation).[3] The transponder model for user ℓ is composed of an IMUX filter which selects the ℓ-th carrier, an HPA, and an OMUX filter which reduces the out-of-band power due to the spectral regrowth after nonlinear amplification [43]. The HPA is a nonlinear memoryless device defined through its AM/AM and AM/PM characteristics, describing the amplitude and phase distortions caused on the signal at its input.

The outputs of different transponders are multiplexed again in the downlink to form the signal $s(t)$, and we assume that the adjacent users have a frequency separation of F_d, usually equal to that in the uplink. Fig. 5.1 shows the system model highlighting the satellite transponder for user with $\ell = 0$. The useful signal at the user terminal is still the sum of independent contributions, one for each transponder (although these contributions are no more, rigorously, linearly modulated due to the nonlinear transformation of the on-board HPA). The received signal is also corrupted by the downlink AWGN, whose low-pass equivalent $w(t)$ has power spectral density (PSD) N_0. The low-pass equivalent of the received signal has thus expression

$$r(t) = s(t) + w(t)\,. \tag{5.2}$$

We remark that, in the simulation results, this system has been simulated with realistic

[3] The multiple-carriers-per-transponder scenario is conceptually similar to that considered in this chapter, the only difference being the fact that more adjacent carriers are amplified by the same HPA. Thus, the effects of nonlinear ICI (intermodulation distortion) become more relevant and proper multicarrier detection or predistortion algorithms could be also considered [56, 57, 58].

assumptions. In other words, the satellite receives the entire signal (5.1). Each carrier is then selected by the IMUX filter of its own transponder,[4] amplified by its own HPA, filtered again by the OMUX filter, and then the signals at the output of all transponders are multiplexed again on air.

We evaluate the ultimate performance limits of this communication system when single-user detection is employed at the receiver side. The proposed technique consists of allowing interference in time and/or frequency by reducing the values of T, F_u, and F_d, (partially) coping with it at the receiver, in order to increase the SE. In other words, T, F_u, and F_d are chosen as the values that give the maximum value of the SE. These values depend on the employed detector—the larger the interference that the receiver can cope with, the larger the SE and the lower the values of time and frequency spacings. Notice that, since we are considering single-user detection, the receiver is not able to deal with the interference due to the overlap of different channels. In this case, the optimization of the frequency spacings is actually an optimization of the frequency guard bands generally introduced in satellite systems to avoid the nonlinear cross-talk, since a single-user receiver can tolerate only a very small amount of ICI.

The considered nonlinear satellite channel reduces, as a particular case, to the linear channel, provided that the HPA is driven far from saturation. Hence, all the considerations in this chapter can be straightforwardly extended to the linear channel case. A few results can be found on Chapter 3.

5.2 Optimization of the spectral efficiency

We describe the framework used to evaluate the ultimate performance limits of the considered satellite system and to perform the optimization of the time and frequency spacings. To simplify the analysis, we will assume $F_u = F_d = F$. We perform this investigation by constraining the complexity of the employed receiver. In particular, as mentioned, we assume that a single-user detector is used. For this reason, without

[4] Being the IMUX a non ideal filter and due to a possible overlap among different carriers, this filtering will not be perfect and ICI will occur.

loss of generality we only consider the detection of symbols $\boldsymbol{c}^{(0)} = \{c_k^{(0)}\}$ of user with $\ell = 0$.[5] In addition, we also consider low-complexity receivers taking into account only a portion of the actual channel memory. Under these constraints, we compute the IR, i.e., the average mutual information when the channel inputs are i.u.d. random variables belonging to a given constellation. Provided that a proper *auxiliary channel* can be defined for which the adopted low-complexity receiver is optimal, the computed IR represents an achievable lower bound of the IR of the actual channel, according to mismatched detection [19].

Denoting by $\boldsymbol{r}^{(0)}$ a set of sufficient statistics for the detection of $\boldsymbol{c}^{(0)}$, the achievable IR, measured in bit per channel use, can be obtained as

$$I_\mathrm{R} = \lim_{N\to\infty} \frac{1}{N} E\left\{ \log \frac{q(\boldsymbol{r}^{(0)}|\boldsymbol{c}^{(0)})}{q(\boldsymbol{r}^{(0)})} \right\}, \tag{5.3}$$

where N is the number of transmitted symbols. The probability density functions $q(\boldsymbol{r}^{(0)}|\boldsymbol{c}^{(0)})$ and $q(\boldsymbol{r}^{(0)})$ are computed by using the optimal maximum-a-posteriori (MAP) symbol detector for the auxiliary channel, while the expectation in (5.3) is with respect to the input and output sequences generated according to the actual channel model [21]. In the next section, we will discuss two different low-complexity detectors for nonlinear satellite channels and we will define the corresponding auxiliary channels.

We can define the user's bandwidth as the frequency separation F between two adjacent carriers. The achievable SE is thus

$$\eta = \frac{I_\mathrm{R}}{FT} \quad \text{[b/s/Hz]}. \tag{5.4}$$

The aim of the proposed technique is to find the values of F and T providing, for each value of the signal-to-noise ratio (SNR), the maximum value of SE achievable by that particular receiver, optimal for the considered specific auxiliary channel. Namely, we compute

$$\eta_\mathrm{M} = \max_{F,T>0} \eta(F,T). \tag{5.5}$$

[5] Assuming a system with an infinite number of users, the results do not depend on a specific user.

Typically, the dependency on the SNR value is not critical, in the sense that we can identify, for each shaping pulse and modulation format, two or at most three SNR regions for which the optimal spacings practically have the same value.

Having removed the constraint of orthogonal signaling, one more degree of freedom in the SE optimization is represented by the bandwidth W of the shaping pulse $p(t)$. Hence, guided by the same idea behind the TF packing technique, we can also optimize W, further increasing both ICI and ISI due to the adjacent users and to the IMUX and OMUX filters, respectively. Whereas on the AWGN channel this optimization is implicit in TF packing, in the sense that we can obtain the same ICI by fixing F and increasing W or by fixing W and decreasing F, this is no more true for our nonlinear channel since IMUX and OMUX bandwidths are kept fixed. Hence, an increased value of W also increases the ISI. The benefit of the bandwidth optimization is twofold: it can be used as an alternative to frequency packing (e.g., in cases where the transponder frequency plan cannot be modified and hence frequency packing is not an option), or it can be used to improve the results of TF packing. In this case, we thus compute

$$\eta_{\mathrm{M}} = \max_{F,T,W>0} \eta(F,T,W)\,. \tag{5.6}$$

For fair comparisons in terms of SE, we need a proper definition of the SNR. We define the SNR as the ratio $P_{\mathrm{sat}}/N_0 F$ between the peak power P_{sat} at the output of an HPA in response to a continuous wave input, denoted as the amplifier saturation power, and the noise power in the bandwidth assigned to each carrier, which coincides with the frequency spacing F between two adjacent carriers. This is because in a satellite forward link, the two main resource constraints are the available frequency spectrum and the radiated power on-board of the satellite. The adopted SNR definition is independent of the transmit waveform and its parameters. This provides a common measure to compare the performance of different solutions in a fair manner. In the following, we also define the output back-off (OBO) for each waveform (or modulation scheme) as the power ratio (in dBs) between the unmodulated carrier at saturation and the modulated carrier after the OMUX.

Without loss of generality, T and F in (5.4)-(5.6) can be normalized to some reference values T_B and F_B. We will denote $\nu = F/F_B$ and $\tau = T/T_B$. In the numerical results, we will choose T_B and F_B as the symbol time and the frequency spacing adopted in the DVB-S2 standard, which is considered as a benchmark scenario.

5.3 Considered detectors and corresponding auxiliary channel models

The system model described in §5.1 is representative of the considered scenario and has been employed in the information-theoretic analysis and in the simulations results. In this section, we describe the employed auxiliary channel models and the corresponding optimal MAP symbol detectors. As explained in §5.2, they are used to compute two lower bounds on the SE for the considered channel [21]. Since these lower bounds are achievable by those receivers, we will say that the computed lower bounds are the SE values of the considered channel when those receivers are employed.

Memoryless model and predistortion at the transmitter

When the HPA AM/AM and AM/PM characteristics are properly estimated and fed back to the transmitter, the sequence of symbols $\{c_k^{(\ell)}\}$ can be properly predistorted to form the sequence $\{c_k^{\prime(\ell)}\}$ that is transmitted instead, in order to compensate for the effect of the non-linearity and possibly to reduce the ISI. Here we consider the dynamic data predistortion technique described in [54, 55] and also suggested for the application in DVB-S2 systems [43], where the symbol $c_k^{\prime(\ell)}$ transmitted by user ℓ at time k is a function of a sequence of $2L_p+1$ input symbols, i.e., $c_k' = f(c_{k-L_p},\ldots,c_k,\ldots,c_{k+L_p})$. The mapping f at the transmitter is implemented through a look-up table (LUT), which is computed through an iterative procedure performed off-line and described in [54, 55] for each modulation format, setting of the system parameters, and SNR value. This procedure searches the best trade-off between the interference reduction and the increase of the OBO. The complexity at the transmitter depends on the number of symbols accounted for through the parameter

L_p. The transmitted signal is thus

$$x(t) = \sum_{\ell} \sum_{k} c_k^{\prime(\ell)} p(t - kT) e^{j2\pi \ell F t} \tag{5.7}$$

whereas, at the receiver, a simple single-user memoryless channel is assumed corresponding to the auxiliary channel (for user with $\ell = 0$)

$$r_k^{(0)} = c_k^{(0)} + n_k \tag{5.8}$$

where n_k is a zero-mean circularly symmetric white Gaussian process with PSD $(N_0 + N_I)$, N_I being a design parameter which can be optimized through computer simulations—an increase of the assumed noise variance can improve the computed achievable lower bound on the SE [5].

Model with memory and advanced detection

A valid alternative to nonlinear compensation techniques at the transmitter relies upon the adoption of advanced detectors which can manage the nonlinear distortions and the ISI. In this work, we consider detection based on an approximate signal model described in [51], which comes from a simplified Volterra series expansion of the nonlinear channel. To limit the receiver complexity with a limited performance degradation, we also apply a CS technique [1]. In fact, when the memory of the channel is too large to be taken into account by a full complexity detector, an excellent performance can be achieved by properly filtering the received signal before adopting a reduced-state detector [1].

A very effective CS technique for general linear channels is described in [2], while its extension to nonlinear satellite systems is reported in Chapter 4. We will denote the memory taken into account by the advanced detection scheme as L_r.

5.4 Numerical results

Spectral efficiency

Considering the DVB-S2 system as a benchmark scenario, we now show the improvement, in terms of SE, that can be obtained by adopting the TF packing technique joint with an advanced processing at the receiver. In the following, we assume as reference the time and frequency spacings of DVB-S2, i.e., $1/T_B = 27.5$ Mbaud and $F_B = 41.5$ MHz, and use these values to normalize all time and frequency spacings. Let us consider a typical DVB-S2 scenario where, at the transmitter, the shaping pulse $p(t)$ has a RRC spectrum, whereas the IMUX and OMUX filters and the nonlinear characteristics of the HPA are those reported in [43, Figs. H.12 and H.13]. The standard considers the following modulation formats: QPSK, 8PSK, 16APSK, and 32APSK. To combat ISI and nonlinear distortions, a data predistorter is employed at the transmitter whereas at the receiver a symbol-by-symbol detector is assumed. Here, we consider the predistorter described in §5.3, with $L_p = 2$ for QPSK, 8PSK 16APSK, and $L_p = 1$ for 32APSK. The SE results have been obtained by computing the IR in (5.3) by means of the Monte Carlo method described in [21]. For each case, i.e., for each modulation format, employed detection algorithm, and choice of the system parameter, we also performed a coarse optimization of the noise variance to be set at receiver [5], and of the amplifier operation point through the OBO. Unless otherwise specified, the roll-off factor of the RRC pulses is $\alpha = 0.2$, which is the lowest value considered in the standard.

We first consider the achievable SE of our benchmark scenario, and in Fig. 5.2 we report η as a function of P_{sat}/N_0F for the four modulation formats of the standard (QPSK, 8PSK, 16APSK, and 32APSK). In this case, it is $W = (1+\alpha)/T_B$. We verified that comparable SE values can be also obtained by using, instead of the predistorter, the advanced detection scheme of §5.3 with $L_r = 0$ (MMSE detection). We also consider two alternative ways that, at least in the case of a linear channel, can be used to improve the SE without resorting to TF packing.[6] The simplest approach relies on the increase of the modulation cardinality, and in Fig. 5.2 we also show the SE for the 64APSK modulation [59]. It can be seen that the 64APSK modulation, due to the higher impact of the non-linearities, allows to increase the SE only at high SNR values and it seems there is no hope to improve the SE in the low and medium SNR regions.

[6] On a nonlinear satellite channel, due to the increased peak-to-average power ratio, their application must be carefully considered since not necessarily produces the expected benefits.

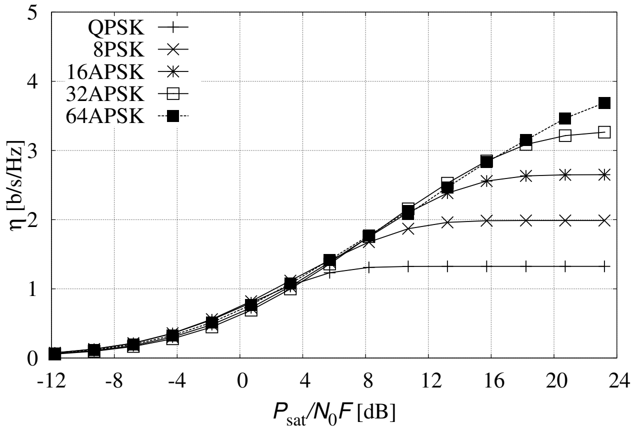

Figure 5.2: Spectral efficiency of DVB-S2 modulations with roll-off 0.2, data predistortion, and memoryless detection. Comparison with a constellation of increased cardinality (64APSK).

An alternative way of improving the SE is based on a reduction of the roll-off factor. In Fig. 5.3, we consider QPSK and 16APSK modulations in a scenario where predistortion at the transmitter and symbol-by-symbol detection at the receiver are still employed. We show the SE improvement that can be obtained by reducing the roll-off to $\alpha = 0.05$.[7] We can observe that the roll-off reduction improves the SE with respect to DVB-S2 for all SNR values. On the other hand, as shown in Fig. 5.3, better results can be obtained by allowing TF packing. The values of T and F are chosen as those providing the largest SE. This search is carried out by evaluating (5.5) on grid of values of T and F (coarse search), followed by interpolation of the obtained values (fine search). We point out that these curves have been obtained without reducing the roll-off factor, which is still $\alpha = 0.2$, and employing the same predistorter and

[7] We properly modify the transmitted signal such that it occupies the same bandwidth as that of the signal with roll-off 0.2. We verified that, in this particular case, no improvement can be obtained by resorting to a more sophisticated receiver based on linear or nonlinear equalization in addition to or in substitution of the predistorter.

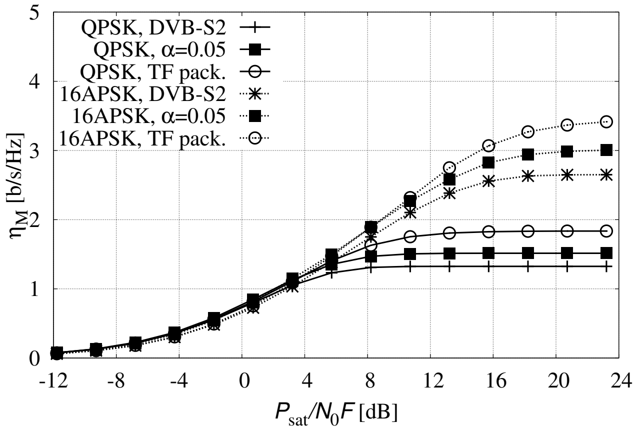

Figure 5.3: Improvements, in terms of spectral efficiency, that can be obtained for QPSK and 16APSK modulations by reducing of the roll-off ($\alpha = 0.05$) or by adopting the TF packing technique. In all cases, predistortion at the transmitter and memoryless detection at the receiver are employed.

symbol-by-symbol receiver adopted in the DVB-S2 system [54, 55].

With the aim of further improving the performance, we now consider TF packing and a system without predistortion at the transmitter but using the advanced detection algorithm described in §5.3, joint with CS ($L_r = 1$, to reduce the receiver complexity, since in this case the BCJR algorithm has only M states). The assumed order of the Volterra model (4.2) is $v = 5$ since, in this case, when considering a single-carrier transmission (no adjacent users) the minimum mean square error between the model and the actual signal is very low. The results for QPSK and 16APSK modulations are reported in Fig. 5.4, where we also show the DVB-S2 benchmark curves discussed above and the curves related to TF packing when predistortion at the transmitter and memoryless detection at the receiver are used. These results show the impressive improvement achievable by TF packing combined with the considered advanced receiver, which, with a memory of only one symbol, can cope with much more interference than the schemes employing the predistorter and a memoryless detector.

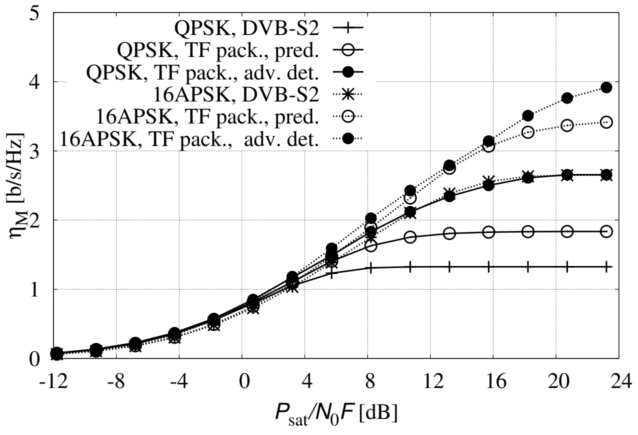

Figure 5.4: Spectral efficiency for QPSK and 16APSK modulations with TF packing and advanced detector (TF pack., adv. det.). Comparison with the DVB-S2 scenario and the case of TF packing when a predistorter and a symbol-by-symbol detector are adopted (TF pack., pred.).

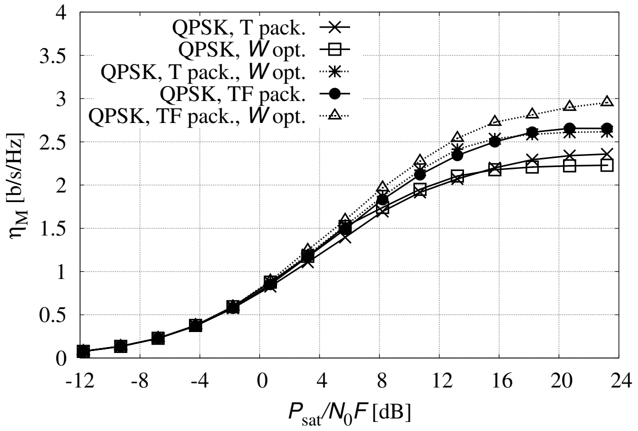

Figure 5.5: Spectral efficiency of QPSK modulation with TF packing and bandwidth optimization by adopting the advanced receiver with CS ($L_r = 1$).

In the previous figures we considered, as mentioned, a bandwidth $W = 1.2/T_B$. We also considered the possibility of optimizing the bandwidth W, as described in §5.2. In Fig. 5.5, we consider QPSK modulation and the advanced receiver with $L_r = 1$. As expected, the combination of TF packing with the bandwidth optimization gives the best results. We also show the results in case only time packing or only the bandwidth optimization are adopted. Interestingly, the SE of time packing with bandwidth optimization is quite similar to that achievable by TF packing.

Finally, to summarize the results, Figure 5.6 shows the SE for all DVB-S2 modulations (QPSK, 8PSK, 16APSK and 32APSK) with TF packing, bandwidth optimization, and the advanced receiver. For clarity, we show only one curve which, for each abscissa, reports only the largest value of the four curves (the "envelope"). In the same figure, we also plot three other SE curves obtained by using predistortion and a memoryless receiver. The lowest one is that corresponding to the DVB-S2 scenario (one curve which is the "envelope" of all four curves in Fig. 5.2), the SE curve for the 64APSK modulation, and the SE curve in case of roll-off $\alpha = 0.05$ reduction. In this latter case, we considered all modulations with cardinality up to 64, and hence this

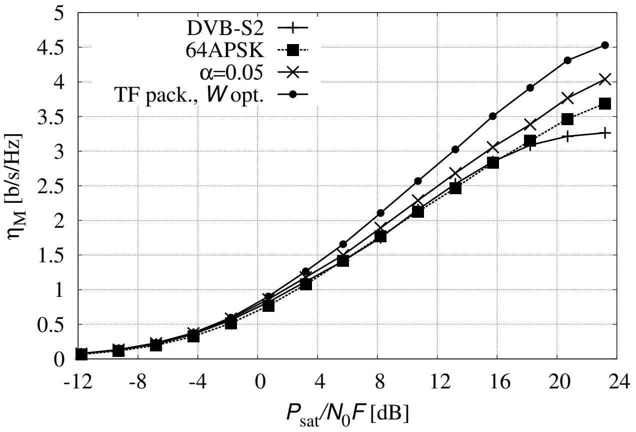

Figure 5.6: Spectral efficiency of TF packing with bandwidth optimization (TF pack., *W* opt.). Comparison with DVB-S2, 64APSK and roll-off reduction.

curve represents the effect of both roll-off reduction and cardinality increase with respect to DVB-S2. The figure shows that TF packing and advanced receiver processing allows a SE improvement of around 40% w.r.t. DVB-S2 at high SNR. This gain is partly due to the fact that current DVB-S2 standard does not support higher order modulations or lower roll-off values. However, there is still considerable SE improvements at lower SNR values.

Modulation and coding formats

What information theory promises can be approached by using proper coding schemes. All the considered modulation and coding formats (MODCODs) use the low-density parity-check (LDPC) codes with length 64800 bits of the DVB-S2 standard. We adopt the optimized values for T, F, and W and the advanced detector described in §5.3. Due to the soft-input soft-output nature of the considered detection algorithm, we can adopt iterative detection and decoding. We distinguish between local iterations, within the LDPC decoder, and global iterations, between the detector and the decoder.

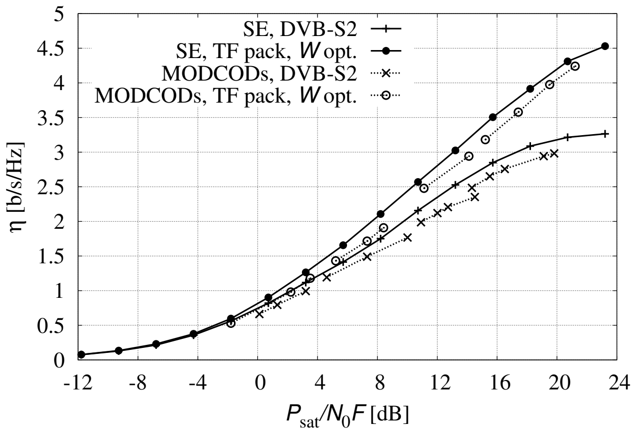

Figure 5.7: Modulation and coding formats of the DVB-S2 standard and comparison with those designed for the proposed TF packing technique with optimized bandwidth.

Here, we allow a maximum of 5 global iterations and 20 local iterations.

BER results have been computed by means of Monte Carlo simulations and are reported in the SE plane in Figure 5.7 using, as reference, a BER of 10^{-6}. In the same figure, the performance of the DVB-S2 MODCODs is also shown for comparison. We recall that for them predistortion at the transmitter and symbol-by-symbol detection at the receiver are adopted. Moreover, for them we have $\tau = 1$ and $\nu = 1$. The details of the considered MODCODs are reported in Tables 5.1 and 5.2. These results are in perfect agreement with the theoretical analysis and confirm that the TF packing technique can provide an impressive performance improvement w.r.t. the DVB-S2 standard.

5.5 Conclusions

We have investigated the TF packing technique, jointly with an advanced processing at the receiver, to improve the spectral efficiency of a nonlinear satellite sys-

Table 5.1: Details of the MODCODs based on TF packing.

	rate	τ	ν	W_{opt}	P/N_0F [dB]	η [b/s/Hz]
QPSK	1/3	0.833	1.00	+20%	-1.7	0.53
	1/2	0.750	0.90	+20%	2.2	0.98
	3/5	0.750	0.90	+20%	3.6	1.18
8PSK	1/2	0.731	0.95	+30%	5.3	1.43
	3/5	0.731	0.95	+30%	7.4	1.72
	2/3	0.731	0.95	+30%	8.5	1.91
16APSK	2/3	0.792	0.90	+20%	11.1	2.48
	3/4	0.750	0.90	+20%	14.1	2.94
32APSK	2/3	0.731	0.95	+30%	15.3	3.18
	3/4	0.731	0.95	+30%	17.5	3.58
	5/6	0.731	0.95	+30%	19.5	3.98
	8/9	0.731	0.95	+30%	21.2	4.24

tem employing linear modulations with finite constellations. As a first step, through an information-theoretic analysis, we computed the spectral efficiency achievable through this technique showing, with reference to the DVB-S2 specifications, that without an advanced processing at the receiver, the potential gains are very limited. On the other hand, a detector which takes into account a memory of only one symbol, and thus with a very limited complexity increase, it is possible to obtain a gain up to 40% in terms of spectral efficiency with respect to the conventional use of the current standard. This impressive gain is partly due to optimized carrier spacing of adjacent transponders. Although this assumption may not be applicable to all satellite communication systems, the results of this chapter indicates possible new system design directions to further improve the spectral efficiency. All these considerations can be extended to other channels and scenarios.

Table 5.2: Details of the DVB-S2 MODCODs. In this case, $\tau = 1$ and $\nu = 1$.

	rate	P/N_0F [dB]	η [b/s/Hz]
QPSK	1/2	0.1	0.66
	3/5	1.4	0.79
	3/4	3.2	0.99
8PSK	3/5	4.6	1.19
	3/4	7.3	1.49
	8/9	10.0	1.77
16APSK	3/4	10.9	1.99
	4/5	12.0	2.12
	5/6	12.7	2.21
	8/9	14.6	2.35
32APSK	3/4	14.3	2.48
	4/5	15.6	2.65
	5/6	16.5	2.76
	8/9	19.2	2.94
	9/10	19.9	2.98

Appendix A

Toeplitz matrix, circulant matrix and Szegö theorem

Toeplitz matrix and circulant matrix are both useful structure to represent channels with memory. In this section we denote a $N \times N$ Toeplitz matrix by $\boldsymbol{T}_N$, where the subscript denotes explicitly its dimension. A Toeplitz matrix has elements $(\boldsymbol{T}_N)_{ij} = t_{i-j}$, being $\{t_i\}$ a sequence, and reads

$$\boldsymbol{T}_N = \begin{pmatrix} t_0 & t_{-1} & \dots & t_{-(N-1)} \\ t_1 & t_0 & \dots & t_{-(N-2)} \\ \vdots & & \ddots & \vdots \\ t_{N-1} & t_{N-2} & \dots & t_0 \end{pmatrix} . \tag{A.1}$$

The most common application of Toeplitz matrix in this thesis, is to represent a filtering. Namely, if we consider a sequence $\{x_i\}$ filtered by $\{t_i\}$, the sequence at the output of the filter reads

$$y_k = \sum_i t_i x_{k-i} . \tag{A.2}$$

The convolution (A.2) can be also written by means of the following matrix notation

$$\boldsymbol{y} = \boldsymbol{T}_N \boldsymbol{x} \tag{A.3}$$

where $\boldsymbol{x} = [x_0, \dots, x_{N-1}]^T$.

A circulant matrix $\boldsymbol{C}_N$ is a special case of Toeplitz matrix constructed on a sequence such that $t_{-k} = t_{N-k}$. With such sequence, the matrix reads

$$\boldsymbol{C}_N = \begin{pmatrix} t_0 & t_{N-1} & t_{N-2} & \cdots & t_1 \\ t_1 & t_0 & t_{N-1} & \cdots & t_2 \\ t_2 & t_1 & t_0 & & t_3 \\ \vdots & & & \ddots & \vdots \\ t_{N-1} & t_{N-2} & \cdots & \cdots & t_0 \end{pmatrix}. \tag{A.4}$$

In other words, in a circulant matrix all rows are cyclic shift of the first row. One important property of a circulant matrix is related to its eigenvalue decomposition [36]. Said $\boldsymbol{F}$ the $N \times N$ Fourier matrix, with elements $(\boldsymbol{F})_{ik} = e^{-j2\pi ik/N}$, it can be shown that any circulant matrix can be decomposed as

$$\boldsymbol{C}_N = \boldsymbol{F}^\dagger \boldsymbol{\Lambda} \boldsymbol{F} \tag{A.5}$$

where $\boldsymbol{\Lambda}$ is the diagonal matrix containing the eigenvalues $(\boldsymbol{\Lambda})_{ii} = C_i$, being C_i the eigenvalues of $\boldsymbol{C}_N$. It can be shown that these eigenvalues are given by the discrete fourier trasform (DFT)[1]

$$C_i = \sum_{k=0}^{N-1} t_k e^{-j2\pi ki/N}. \tag{A.6}$$

From the two definitions of Toeplitz matrix (A.1) and circulant matrix (A.4) it can be expected that Toeplitz matrix and circulant matrix can have similar properties. In particular, by using naïve words, we can expect that a Toeplitz matrix $\boldsymbol{T}_N$, for $N \to \infty$ *behaves* like a circulant matrix, and thus also its eigenvalues are related to the Fourier transform of $\{t_i\}$. More formally this relation is given by the Szegö theorem.

Theorem 3 (Szegö theorem)**.** *Let $\{\boldsymbol{T}_N\}$ be a sequence of $N \times N$ Toeplitz matrix such that $\{t_i\}$ is absolutely summable. Let $\{\tau_{N,i}\}_{i=0}^{N-1}$ be the eigenvalues of $\boldsymbol{T}_N$ and s any positive definite integer. Then*

$$\lim_{N\to\infty} \frac{1}{N} \sum_{i=0}^{N-1} \tau_{N,i}^s = \frac{1}{2\pi} \int_0^{2\pi} T^s(\omega) \mathrm{d}\omega \tag{A.7}$$

[1] The DFT must not be confused with the discrete time Fourier transform (DTFT).

where $T(\omega)$ is the discrete time Fourier transform (DTFT)

$$T(\omega) = \sum_{i=0}^{N-1} t_i e^{-j\omega i}. \tag{A.8}$$

Furthermore, if $T(\omega)$ is real, with essential infimum m and essential supremum M, then for any continuous function $f : [m, M] \to [0, \infty)$

$$\lim_{N\to\infty} \frac{1}{N} \sum_{i=0}^{N-1} f(\tau_{N,i}) = \frac{1}{2\pi} \int_0^{2\pi} f(T(\omega)) \mathrm{d}\omega. \tag{A.9}$$

A simple application of the theorem, is given by setting f as the logarithm. This gives the identity

$$\lim_{N\to\infty} \frac{1}{N} \log\det \boldsymbol{T}_N = \frac{1}{2\pi} \int_0^{2\pi} \log(T(\omega)) \mathrm{d}\omega. \tag{A.10}$$

If we now consider as example the matrix channel (1.8), its capacity is given by $\log_2 \det(\boldsymbol{I} + \boldsymbol{G}/N_0)$, which for $N \to \infty$ tends to

$$\lim_{N\to\infty} \frac{1}{N} \log_2 \det\left(\boldsymbol{I} + \frac{\boldsymbol{G}}{N_0}\right) = \frac{1}{2\pi} \int_0^{2\pi} \log\left(1 + \frac{G(\omega)}{N_0}\right) \mathrm{d}\omega \tag{A.11}$$

as shown in [35].

For further details with a very clean explation, further consequences and applications, the reader can see [60].

Appendix B

CS for channels represented by a block Toeplitz matrix

In this appendix we derive the CS solution when the channel model reads

$$\mathbf{r} = \mathbf{H}\mathbf{c} + \mathbf{w}\,, \tag{B.1}$$

where $\mathbf{H}$ is block lower triangular and Toeplitz matrix with size $KN \times KN$ built from a sequence of matrices $\{\boldsymbol{H}_i\}_{i \geq 0}$ with size $K \times K$. Namely it reads

$$\mathbf{H} = \begin{pmatrix} \boldsymbol{H}_0 & \mathbf{0} & \dots & \mathbf{0} \\ \boldsymbol{H}_1 & \boldsymbol{H}_0 & \dots & \mathbf{0} \\ \vdots & & \ddots & \vdots \\ \boldsymbol{H}_{N-1} & \boldsymbol{H}_{N-2} & \dots & \boldsymbol{H}_0 \end{pmatrix}. \tag{B.2}$$

$\mathbf{c}$ is assumed to be a block vector of complex Gaussian random variables, with mean zero and autocorrelation matrix $\mathbf{V} = \mathrm{E}\{\mathbf{c}\mathbf{c}^\dagger\}$. We constrain the autocorrelation matrix $\mathbf{V}$ to be block diagonal as

$$\mathbf{V} = \begin{pmatrix} \boldsymbol{V} & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \boldsymbol{V} & \dots & \mathbf{0} \\ \vdots & & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \dots & \boldsymbol{V} \end{pmatrix}, \tag{B.3}$$

where $\boldsymbol{V}$ is a positive definite matrix. The CS detector considers a mismatched channel law

$$q(\mathbf{r}|\mathbf{c}) \propto \exp\left\{\Re\left(\mathbf{c}^\dagger(\mathbf{H}^r)^\dagger\mathbf{r}\right) - \mathbf{c}^\dagger\mathbf{G}^r\mathbf{c}\right\} . \tag{B.4}$$

where the channel shortener $\mathbf{H}^r$ and the target response $\mathbf{G}^r$ are block Toeplitz matrices such that the AIR is maximized for a given memory L taken into account at the detector. The target response has constraint

$$(\mathbf{G}^r)_{ij} = \mathbf{0} \quad \forall|i-j| > L \tag{B.5}$$

where with $(\mathbf{G}^r)_{ij}$ we mean the (i,j) block, and $\mathbf{0}$ is $K\times K$ matrix of all zeros.

To derive the optimal channel shortener and target response, as first step we need a closed formula for the AIR

$$\begin{aligned} I_\mathrm{R} &= \mathfrak{h}(\boldsymbol{r}) - \mathfrak{h}(\boldsymbol{r}|\boldsymbol{c}) & \text{(B.6)} \\ &= \mathrm{E}\left\{\log_2\frac{q(\boldsymbol{r}|\boldsymbol{c})}{q(\boldsymbol{r})}\right\} . & \text{(B.7)} \end{aligned}$$

$q(\boldsymbol{r})$ is found to be

$$\begin{aligned} q(\mathbf{r}) &= \frac{1}{\pi^{KN}\det(\mathbf{V})}\int q\left(\mathbf{r}|\mathbf{c}\right)\exp\left\{-\mathbf{c}^\dagger\mathbf{V}^{-1}\mathbf{c}\right\}\mathrm{d}\mathbf{c} & \text{(B.8)} \\ &= \frac{1}{\det(\mathbf{G}^r\mathbf{V}+\mathbf{I})}\exp\{\mathbf{d}^\dagger\left(\mathbf{G}^r+\mathbf{V}^{-1}\right)^{-1}\mathbf{d}\} & \text{(B.9)} \end{aligned}$$

where $\mathbf{d} = \mathbf{H}^r\mathbf{r}$. Therefore,

$$\mathfrak{h}(\boldsymbol{r}) = \log\det(\mathbf{G}^r\mathbf{V}+\mathbf{I}) - \mathrm{Tr}\left((\mathbf{H}^r)^\dagger\left[\mathbf{H}\mathbf{V}\mathbf{H}^\dagger+N_0\mathbf{I}\right]\mathbf{H}^r(\mathbf{G}^r+\mathbf{V}^{-1})^{-1}\right) \tag{B.10}$$

and

$$\mathfrak{h}(\boldsymbol{r}|\boldsymbol{c}) = \mathrm{Tr}(\mathbf{G}^r\mathbf{V}) - 2\Re\left(\mathrm{Tr}\left((\mathbf{H}^r)^\dagger\mathbf{H}\mathbf{V}\right)\right) . \tag{B.11}$$

The derivative of I_R w.r.t. $(\mathbf{H}^r)^\dagger$ is

$$\frac{\partial I_\mathrm{R}}{\partial(\mathbf{H}^r)^\dagger} = (\mathbf{H}\mathbf{V})^T - \left(\left[\mathbf{H}\mathbf{V}\mathbf{H}^\dagger+N_0\mathbf{I}\right]\mathbf{H}^r(\mathbf{G}^r+\mathbf{V}^{-1})^{-1}\right)^T . \tag{B.12}$$

By setting the derivative to zero, we obtain that the optimal filter $\mathbf{H}^r$ is

$$\mathbf{H}^r = \left[\mathbf{H}\mathbf{V}\mathbf{H}^\dagger+N_0\mathbf{I}\right]^{-1}\mathbf{H}\mathbf{V}\left(\mathbf{G}^r+\mathbf{V}^{-1}\right) . \tag{B.13}$$

Using (B.13), the I_{R} is

$$I_{\mathrm{R}} = \frac{1}{N}\left(\log(\det(\mathbf{U}^{\dagger}\mathbf{U}\mathbf{V})) - \mathrm{Tr}\left(\mathbf{U}\mathbf{B}\mathbf{U}^{\dagger}\right) + KN\right) \tag{B.14}$$

where $\mathbf{U}$ is obtained from the Cholesky decomposition $\mathbf{G}^{r} + \mathbf{V}^{-1} = \mathbf{U}^{\dagger}\mathbf{U}$ and

$$\mathbf{B} = \mathbf{V} - \mathbf{V}\mathbf{H}^{\dagger}\left[\mathbf{H}\mathbf{V}\mathbf{H}^{\dagger} + N_0\mathbf{I}\right]^{-1}\mathbf{H}\mathbf{V}\,. \tag{B.15}$$

The (m,n) entries of $\mathbf{U}$ and $\mathbf{B}$ will be denoted by $\boldsymbol{U}_{mn}$ and $\boldsymbol{B}_{mn}$, respectively. Since $\det(\mathbf{U}^{\dagger}\mathbf{U}\mathbf{V})$ depends only on the diagonal elements of $\mathbf{U}$, we can optimize I_{R} over the diagonal of $\mathbf{U}$ and the off-diagonal elements separately. We define $\underline{\mathbf{U}}_n = [\boldsymbol{U}_{n\,n+1}, .., \boldsymbol{U}_{n\,\min(n+L,N)}]$, $\underline{\mathbf{B}}_n = [\boldsymbol{B}_{n\,n+1}, .., \boldsymbol{B}_{n\,\min(n+L,N)}]$,

$$\mathbf{B}_n = \begin{bmatrix} \boldsymbol{B}_{(n+1)\,(n+1)} & \cdots & \boldsymbol{B}_{\min(n+1,L)\,(n+L)} \\ \vdots & \ddots & \vdots \\ \boldsymbol{B}_{\min(n+L,N)\,(n+1)} & \cdots & \boldsymbol{B}_{\min(n+1,L)\,\min(n+1,L)} \end{bmatrix} \tag{B.16}$$

and finally

$$\boldsymbol{\mathcal{C}}_n = \boldsymbol{B}_{nn} - \underline{\mathbf{B}}_n\mathbf{B}_n^{-1}(\underline{\mathbf{B}}_n)^{\dagger}\,. \tag{B.17}$$

Now the trace $\mathrm{Tr}\left(\mathbf{U}\mathbf{B}\mathbf{U}^{\dagger}\right)$ can be rewritten as

$$\sum_n \mathrm{Tr}\left([\boldsymbol{U}_{nn}\,\underline{\mathbf{U}}_n]\begin{bmatrix} \boldsymbol{B}_{nn} & \underline{\mathbf{B}}_n \\ \underline{\mathbf{B}}_n^{\dagger} & \mathbf{B}_n \end{bmatrix}\begin{bmatrix} \boldsymbol{U}_{nn}^{\dagger} \\ \underline{\mathbf{U}}_n^{\dagger} \end{bmatrix}\right)\,. \tag{B.18}$$

Setting its derivative w.r.t. $\underline{\mathbf{U}}_n^{\dagger}$ to zero gives

$$\frac{\partial}{\partial\underline{\mathbf{U}}_n^{\dagger}}\mathrm{Tr}\left(\mathbf{U}\mathbf{B}\mathbf{U}^{\dagger}\right) = (\boldsymbol{U}_{nn}\underline{\mathbf{B}}_n)^{T} + (\underline{\mathbf{U}}_n\mathbf{B}_n)^{T} = \mathbf{0} \tag{B.19}$$

which gives

$$\underline{\mathbf{U}}_n = -\boldsymbol{U}_{nn}\underline{\mathbf{B}}_n\mathbf{B}_n^{-1}\,. \tag{B.20}$$

Replacing (B.20) in (B.14) we find

$$I_{\mathrm{R}} = \frac{1}{N}\log\det(\mathbf{V}) + K + \frac{1}{N}\sum_n \log(\det(\boldsymbol{U}_{nn}^{\dagger}\boldsymbol{U}_{nn})) - \mathrm{Tr}\left(\boldsymbol{U}_{nn}\boldsymbol{\mathcal{C}}_n\boldsymbol{U}_{nn}^{\dagger}\right) \tag{B.21}$$

that can be maximized by setting its derivative w.r.t. $\boldsymbol{U}_{nn}^{\dagger}$ equal to zero. This gives that

$$\frac{\partial I_{\mathrm{R}}}{\partial \boldsymbol{U}_{nn}^{\dagger}} = (\boldsymbol{U}_{nn}^{*})^{-1} - (\boldsymbol{U}_{nn}\boldsymbol{\mathcal{C}}_n)^T = \mathbf{0} \tag{B.22}$$

and the optimal $\boldsymbol{U}_{nn}$ is given by the Cholesky decomposition

$$\boldsymbol{\mathcal{C}}_n^{-1} = \boldsymbol{U}_{nn}^{\dagger}\boldsymbol{U}_{nn} \tag{B.23}$$

Inserting (B.23) into (B.21), the AIR for Gaussian symbols is

$$I_{\mathrm{R}} = \frac{1}{N}\log\det(\mathbf{V}) + \frac{1}{N}\sum_{n}\log(\det(\boldsymbol{\mathcal{C}}_n^{-1}))\,. \tag{B.24}$$

When $N \to \infty$, all $\boldsymbol{\mathcal{C}}_n^{-1}$ are the same, and we obtain the stationary solution

$$I_{\mathrm{R}} = \log\det(\boldsymbol{V}) + \log\det(\boldsymbol{\mathcal{C}}^{-1}) \tag{B.25}$$

and (B.13), (B.15) become stationary as

$$\begin{aligned} \boldsymbol{H}^r(\omega) &= \left[\boldsymbol{H}(\omega)\boldsymbol{V}\boldsymbol{H}^{\dagger}(\omega) + N_0\boldsymbol{I}\right]^{-1}\boldsymbol{H}(\omega)\boldsymbol{V}\left(\boldsymbol{G}^r(\omega) + \boldsymbol{V}^{-1}\right) && \text{(B.26)} \\ \boldsymbol{B}(\omega) &= \boldsymbol{V} - \boldsymbol{V}\boldsymbol{H}^{\dagger}(\omega)\left[\boldsymbol{H}(\omega)\boldsymbol{V}\boldsymbol{H}^{\dagger}(\omega) + N_0\boldsymbol{I}\right]^{-1}\boldsymbol{H}(\omega)\boldsymbol{V}\,. && \text{(B.27)} \end{aligned}$$

Appendix C

Proof of Theorem 1

We first note that $P(\omega)$ only enters the optimization through its square magnitude, and we therefore make the variable substitution $S_p(\omega) = |P(\omega)|^2$ and optimize over $S_p(\omega)$ instead.

The proof will consist of three steps

- A formula for stationary points.
- The observation that some of these do not have strictly positive spectrum.
- Fixing the problem identified in the previous bullets.

Let us now start with the first bullet. From Cramer's rule [36], we get that

$$\boldsymbol{B}^{-1} = \frac{1}{\det(\boldsymbol{B})}[C_{ij}],$$

where C_{ij} is the cofactor of entry (i,j) in $\boldsymbol{B}$. This implies that in (2.8) we can express $\boldsymbol{b}\boldsymbol{B}^{-1}\boldsymbol{b}^\dagger$ as

$$\frac{\sum_{m=1}^{M} \alpha_m b_0^{\phi_{m,0}} b_1^{\phi_{m,1}} (b_1^*)^{\phi_{m,2}} \cdots b_L^{\phi_{m,2L-1}} (b_L^*)^{\phi_{m,2L}}}{\sum_{n=1}^{N} \beta_n b_0^{\psi_{n,0}} b_1^{\psi_{n,1}} (b_1^*)^{\phi_{m,2}} \cdots b_{L-1}^{\psi_{n,2L-3}} (b_{L-1}^*)^{\psi_{n,2L-2}}},$$

where M and N are finite constants that depend on L, $\alpha_m, \beta_m \in \{\pm 1\}$, and both $\phi_{m,\ell}$ and $\psi_{n,\ell}$ are non-negative integers which satisfy

$$\sum_{\ell=0}^{2L} \phi_{m,\ell} = L+1 \quad \text{and} \quad \sum_{\ell=0}^{2L-2} \psi_{n,\ell} = L.$$

We next introduce the variable substitution

$$y(\omega)=\frac{N_0}{|H(\omega)|^2S_p(\omega)+N_0},\ S_p(\omega)=\frac{N_0}{|H(\omega)|^2}\left[\frac{1}{y(\omega)}-1\right].$$

The constraint $\int S_p(\omega)\mathrm{d}\omega=2\pi$ translates into

$$e[y(\omega)]\triangleq\int_{-\pi}^{\pi}\frac{1}{y(\omega)|H(\omega)|^2}\mathrm{d}\omega=\int_{-\pi}^{\pi}\frac{1}{|H(\omega)|^2}\mathrm{d}\omega+\frac{2\pi}{N_0}.$$

Furthermore, we have

$$b_i=\frac{1}{2\pi}\int_{-\pi}^{\pi}y(\omega)e^{j\omega i}\mathrm{d}\omega.$$

The constrained Euler-Lagrange equation [61] becomes

$$\frac{\delta\mathscr{C}}{\delta y}=\lambda\frac{\delta e}{\delta y}=-\frac{\lambda}{|H(\omega)|^2y^2(\omega)}.$$

The functional derivative $\delta b_k^s/\delta y$ equals

$$\begin{aligned}\frac{\delta b_i^s}{\delta y} &= \frac{\delta\left[\int_{-\pi}^{\pi}y(\omega)e^{j\omega i}\mathrm{d}\omega\right]^s}{\delta y}\\ &= s\left[\int_{-\pi}^{\pi}y(\omega)e^{j\omega i}\mathrm{d}\omega\right]^{s-1}e^{j\omega i}\\ &= sb_i^{s-1}e^{j\omega i}.\end{aligned}$$

We now note that b_i, raised to any power, is a *constant* that depends explicitly on y. Therefore, by an application on the quotient rule for the derivative and the chain rule to (2.8), we obtain an expression of the form

$$\frac{\delta\mathscr{C}}{\delta y}=1-\frac{\sum_{\ell=-L}^{L}A_\ell[y(\omega)]e^{j\ell\omega}}{C[y(\omega)]},$$

where the constants $A_\ell[y(\omega)]$ and $C[y(\omega)]$ explicitly depend on $y(\omega)$, e.g.,

$$C[y(\omega)]=\left[\sum_{n=1}^{N}\beta_nb_0^{\psi_{n,0}}b_1^{\psi_{n,1}}\cdots b_{L-1}^{\psi_{n,2L-3}}(b_{L-1}^*)^{\psi_{n,2L-2}}\right]^2.$$

By manipulation of the Euler-Lagrange equation and by introducing a new set of constants $\{B_\ell[y(\omega)]\}$, we obtain

$$y(\omega) = \frac{1}{\sqrt{|H(\omega)|^2[\sum_{\ell=-L}^{L} B_\ell[y(\omega)]e^{j\ell\omega}]}}.$$

This translates into a general form of the optimal $S_p(\omega)$ which reads

$$S_p^{\mathrm{opt}}(\omega) = \frac{N_0}{\sqrt{|H(\omega)|^2}} \sqrt{\sum_{\ell=-L}^{L} A_\ell e^{j\ell\omega}} - \frac{N_0}{|H(\omega)|^2} \tag{C.1}$$

where coefficients A_ℓ must have a Hermitian symmetry.

We have now found a general form for any stationary point. Unfortunately, for a given $H(\omega)$, this stationary point may lie outside the domain of the optimization. The optimal spectrum $S_p(\omega)$ must therefore lie on the boundary of the optimization domain, which in this case implies that $S_p(\omega) = 0$ for $\omega \in \mathscr{I}_0 \subset [-\pi, \pi]$. Let us define $\mathscr{I}_+$ as the subset $[-\pi, \pi]$ where $S_p(\omega) > 0$ except for the endpoints of $\mathscr{I}_+$ where $S_p(\omega) = 0$ due to the assumption of a continuous spectrum. Note that $\mathscr{I}_+$ may be the union of several disjoint sub-intervals of $[-\pi, \pi]$. We can now rewrite the constraint and the expressions of b_k as

$$e[y(\omega)] = \int_{\mathscr{I}_+} \frac{1}{|H(\omega)|^2} \mathrm{d}\omega + \frac{2\pi}{N_0}$$

and

$$b_i = \frac{1}{2\pi} \int_{\mathscr{I}_+} y(\omega) e^{j\omega i} \mathrm{d}\omega.$$

From the first part of the proof, i.e., identifying a necessary condition for stationary points, we have that (C.1) must hold within the interval $\mathscr{I}_+$, and the constants $\{A_\ell\}$ must be such that $S_p^{\mathrm{opt}}(\omega) = 0$ at the end-points of each sub-interval within $\mathscr{I}_+$. Hence, no matter what $\mathscr{I}_+$ is, we can express the optimal $S_p^{\mathrm{opt}}(\omega)$ as in (2.27).

Appendix D

Proof of theorem 2

The waterfilling algorithm provides a transmit filter that satisfies [35]

$$|P(\omega)|^2 = \max\left(0, \theta - \frac{N_0}{|H(\omega)|^2}\right) \tag{D.1}$$

for some power constant θ. In view of Theorem 1, $|P(\omega)|^2$ in (D.1) must also satisfy (2.27). Equating (D.1) and (2.27) yields

$$\theta - \frac{N_0}{|H(\omega)|^2} = \frac{N_0}{\sqrt{|H(\omega)|^2}} \sqrt{\sum_{\ell=-v_C}^{v_C} A_\ell e^{j\ell\omega}} - \frac{N_0}{|H(\omega)|^2}. \tag{D.2}$$

From (D.2), it can be seen that we must have

$$\sum_{\ell=-v_C}^{v_C} A_\ell e^{j\ell\omega} = \gamma |H(\omega)|^2, \tag{D.3}$$

for some constant γ. However,

$$|H(\omega)|^2 = \left|\sum_{\ell=0}^{v} h_\ell e^{-j\ell\omega}\right|^2 = \sum_{\ell=-v}^{v} g_\ell e^{-j\ell\omega}, \tag{D.4}$$

where

$$g_\ell = \sum_k h_k h^*_{k-\ell}. \tag{D.5}$$

Clearly, to satisfy

$$\sum_{\ell=-\nu_C}^{\nu_C} A_\ell e^{j\ell\omega} = \gamma \left[\sum_{\ell=-\nu}^{\nu} g_\ell e^{-j\ell\omega} \right],$$

ν_C must at least equal ν.

Bibliography

[1] D. D. Falconer and F. Magee, “Adaptive channel memory truncation for maximum likelihood sequence estimation,” *Bell System Tech. J.*, vol. 52, pp. 1541–1562, Nov. 1973.

[2] F. Rusek and A. Prlja, “Optimal channel shortening for MIMO and ISI channels,” *IEEE Trans. Wireless Commun.*, vol. 11, pp. 810–818, Feb. 2012.

[3] J. E. Mazo, “Faster-than-Nyquist signaling,” *Bell System Tech. J.*, vol. 54, pp. 1450–1462, Oct. 1975.

[4] J. E. Mazo and H. J. Landau, “On the minimum distance problem for faster-than-Nyquist signaling,” *IEEE Trans. Inform. Theory*, pp. 1420–1427, Nov. 1988.

[5] A. Barbieri, D. Fertonani, and G. Colavolpe, “Time-frequency packing for linear modulations: spectral efficiency and practical detection schemes,” *IEEE Trans. Commun.*, vol. 57, pp. 2951–2959, Oct. 2009.

[6] G. Colavolpe, T. Foggi, A. Modenini, and A. Piemontese, “Faster-than-Nyquist and beyond: how to improve spectral efficiency by accepting interference,” *Opt. Express*, vol. 19, pp. 26600–26609, Dec 2011.

[7] G. Colavolpe and T. Foggi, “High spectral efficiency for long-haul optical links: time-frequency packing vs high-order constellations,” in *Proc. European Conf. on Optical Commun. (ECOC)*, (London, UK), Sept. 2013.

[8] G. Colavolpe and T. Foggi, “Next-generation long-haul optical links: higher spectral efficiency through time-frequency packing,” in *Proc. IEEE Global Telecommun. Conf.*, (Atlanta, U.S.A.), Dec. 2013.

[9] G. Ungerboeck, “Adaptive maximum likelihood receiver for carrier-modulated data-transmission systems,” *IEEE Trans. Commun.*, vol. com-22, pp. 624–636, May 1974.

[10] G. D. Forney, Jr., “Maximum-likelihood sequence estimation of digital sequences in the presence of intersymbol interference,” *IEEE Trans. Inform. Theory*, vol. 18, pp. 284–287, May 1972.

[11] H. Meyr, M. Oerder, and A. Polydoros, “On sampling rate, analog prefiltering, and sufficient statistics for digital receivers,” *IEEE Trans. Commun.*, vol. 42, pp. 3208–3214, Dec. 1994.

[12] L. R. Bahl, J. Cocke, F. Jelinek, and J. Raviv, “Optimal decoding of linear codes for minimizing symbol error rate,” *IEEE Trans. Inform. Theory*, vol. 20, pp. 284–287, Mar. 1974.

[13] F. R. Kschischang, B. J. Frey, and H.-A. Loeliger, “Factor graphs and the sum-product algorithm,” *IEEE Trans. Inform. Theory*, vol. 47, pp. 498–519, Feb. 2001.

[14] P. Roberston, E. Villebrun, and P. Hoeher, “Optimal and sub-optimal maximum a posteriori algorithms suitable for turbo decoding,” *European Trans. Telecommun.*, vol. 8, pp. 119–125, March/April 1997.

[15] D. Fertonani, A. Barbieri, and G. Colavolpe, “Reduced-complexity BCJR algorithm for turbo equalization,” *IEEE Trans. Commun.*, vol. 55, pp. 2279–2287, Dec. 2007.

[16] G. Colavolpe and A. Barbieri, “On MAP symbol detection for ISI channels using the Ungerboeck observation model,” *IEEE Commun. Letters*, vol. 9, pp. 720–722, Aug. 2005.

[17] A. Piemontese, N. Mazzali, and G. Colavolpe, “Improving the spectral efficiency of FDM-CPM systems through packing and multiuser processing,” *International Journal of Satellite Communications and Networking*, vol. 30, pp. 62 – 72, Feb. 2012.

[18] C. Shannon, “A mathematical theory of communication,” *Bell System Tech. J.*, pp. 379–423, July 1948.

[19] N. Merhav, G. Kaplan, A. Lapidoth, and S. Shamai, “On information rates for mismatched decoders,” *IEEE Trans. Inform. Theory*, vol. 40, pp. 1953–1967, Nov. 1994.

[20] A. Ganti, A. Lapidoth, and I. E. Telatar, “Mismatched decoding revisited: General alphabets, channels with memory, and the wide-band limit,” *IEEE Trans. Inform. Theory*, vol. 46, pp. 2315–2328, Nov. 2000.

[21] D. M. Arnold, H.-A. Loeliger, P. O. Vontobel, A. Kavčić, and W. Zeng, “Simulation-based computation of information rates for channels with memory,” *IEEE Trans. Inform. Theory*, vol. 52, pp. 3498–3508, Aug. 2006.

[22] A. Prlja and J. B. Anderson, “Reduced-complexity receivers for strongly narrowband intersymbol interference introduced by faster-than-Nyquist signaling,” *IEEE Trans. Commun.*, vol. 60, no. 9, pp. 2591–2601, 2012.

[23] J. Boutros, N. Gressety, L. Brunel, and M. Fossorier, “Soft-input soft-output lattice sphere decoder for linear channels,” in *Proc. IEEE Global Telecommun. Conf.*, Dec. 2003.

[24] F. Rusek and D. Fertonani, “Bounds on the information rate of intersymbol interference channels based on mismatched receivers,” *IEEE Trans. Inform. Theory*, vol. 58, pp. 1470–1482, Mar. 2012.

[25] G. Colavolpe, A. Modenini, and F. Rusek, “Channel shortening for nonlinear satellite channels,” *IEEE Commun. Letters*, vol. 16, pp. 1929–1932, Dec. 2012.

[26] A. Modenini, F. Rusek, and G. Colavolpe, "Optimal transmit filters for constrained complexity channel shortening detectors," in *Proc. IEEE Intern. Conf. Commun.*, (Budapest, Hungary), pp. 1688–1693, June 2013.

[27] N. Al-Dhahir and J. M. Cioffi, "Efficiently computed reduced-parameter input-aided MMSE equalizers for ML detection: A unified approach," *IEEE Trans. Inform. Theory*, vol. 42, pp. 903–915, Apr. 1996.

[28] N. Al-Dhahir, "FIR channel-shortening equalizers for MIMO ISI channels," *IEEE Trans. Commun.*, vol. 49, pp. 213–218, Feb. 2001.

[29] A. Prlja, *Reduced Receivers for Faster-than-Nyquist Signaling and General Linear Channels*. PhD thesis, Lund University, Lund, Sweden, 2013.

[30] P. J. W. Melsa, R. C. Younce, and C. E. Rohrs, "Impulse response shortening for discrete multitone transceivers," *IEEE Trans. Commun.*, vol. 44, pp. 1662–1672, Dec. 1996.

[31] J. Balakrishnan, R. Martin, and C. Johnson, "Blind, adaptive channel shortening by sum-squared auto-correlation minimization (sam)," *IEEE Trans. Signal Processing*, vol. 51, pp. 3086–3093, Dec. 2003.

[32] I. Abou-Faycal and A. Lapidoth, "On the capacity of reduced complexity receivers for intersymbol interference channels," in *Proc. Conf. on Inform. Sciences and Systems*, (Princeton University,USA), pp. 32–37, Mar. 2000.

[33] J. M. Cioffi, *Data tranmission Theory*. Course text for EE379A-B, chapter 3, http://www.stanford.edu/group/cioffi/.

[34] S. Haykin, *Adaptive Filter Theory*. Englewood Cliffs, NJ: Prentice-Hall, 3rd ed., 1996.

[35] W. Hirt, *Capacity and information rates of discrete-time channels with memory*. PhD thesis, Inst. Signal and Information Processing, Swiss Federal Institute of Technology, Zurich, 1988.

[36] R. A. Horn and C. R. Johnson, *Matrix Analysis*. New York, U.S.A.: Cambridge University Press, 1985.

[37] E. Telatar, "Capacity of multi-antenna Gaussian channels," *European Trans. Telecommun.*, vol. 10, no. 6, pp. 585–596, 1999.

[38] A. Liveris and C. N. Georghiades, "Exploiting faster-than-Nyquist signaling," *IEEE Trans. Commun.*, vol. 47, pp. 1502–1511, Sept. 2003.

[39] F. Rusek and J. B. Anderson, "The two dimensional Mazo limit," in *Proc. IEEE International Symposium on Information Theory*, (Adelaide, Australia), pp. 970–974, Nov. 2005.

[40] A. Modenini, G. Colavolpe, and N. Alagha, "How to significantly improve the spectral efficiency of linear modulations through time-frequency packing and advanced processing," in *Proc. IEEE Intern. Conf. Commun.*, (Ottawa, Canada), pp. 3299–3304, June 2012.

[41] F. Rusek and J. B. Anderson, "Constrained capacities for faster-than-Nyquist signaling," *IEEE Trans. Inform. Theory*, vol. 55, pp. 764 –775, Feb. 2009.

[42] A. V. Oppenheim and R. W. Schafer, *Discrete-Time Signal Processing*. Englewood Cliffs, New Jersey: Prentice-Hall, 1989.

[43] ETSI EN 301 307 Digital Video Broadcasting (DVB); V1.1.2 (2006-06), Second generation framing structure, channel coding and modulation systems for Broadcasting, Interactive Services, News Gathering and other Broadband satellite applications, 2006. Available on ETSI web site (http://www.etsi.org).

[44] F. Rusek and J. B. Anderson, "On information rates of faster than Nyquist signaling," in *Proc. IEEE Global Telecommun. Conf.*, (San Francisco, CA, U.S.A.), Nov. 2006.

[45] F. Rusek and J. B. Anderson, "Maximal capacity partial response signaling," in *Proc. IEEE Intern. Conf. Commun.*, (Glasgow, Scotland), pp. 821–826, June 2007.

[46] A. R. Calderbank and L. H. Ozarow, "Nonequiprobable signalling on the Gaussian channel," *IEEE Trans. Inform. Theory*, vol. 36, pp. 726–740, July 1990.

[47] F. Rusek and D. Fertonani, "Lower bounds on the information rate of intersymbol interference channels based on the ungerboeck observation model," in *Proc. IEEE International Symposium on Information Theory*, 2009.

[48] G. Colavolpe, D. Fertonani, and A. Piemontese, "SISO detection over linear channels with linear complexity in the number of interferers," *IEEE J. of Sel. Topics in Signal Proc.*, vol. 5, pp. 1475–1485, Dec. 2011.

[49] D. Slepian and H. O. Pollak, "Prolate spheroidal wave functions, Fourier analysis and uncertainty - I," *Bell System Tech. J.*, vol. 40, pp. 43–63, Jan. 1961.

[50] A. N. D'Andrea, V. Lottici, and R. Reggiannini, "RF power amplifier linearization through amplitude and phase predistortion," *IEEE Trans. Commun.*, vol. 44, pp. 1477–1484, Nov 1996.

[51] G. Colavolpe and A. Piemontese, "Novel SISO detection algorithms for nonlinear satellite channels," *IEEE Wireless Commun. Letters*, vol. 1, pp. 22–25, Feb. 2012.

[52] S. Benedetto and E. Biglieri, "Nonlinear equalization of digital satellite channels," *IEEE J. Select. Areas Commun.*, vol. 1, pp. 57–62, Jan. 1983.

[53] A. A. M. Saleh, "Frequency-independent and frequency-dependent nonlinear models of TWT amplifiers," *IEEE Trans. Commun.*, vol. 29, pp. 1715–1720, Nov. 1981.

[54] G. Karam and H. Sari, "A data predistortion technique with memory for QAM radio systems," *IEEE Trans. Commun.*, vol. 39, pp. 336–344, Feb. 1991.

[55] E. Casini, R. De Gaudenzi, and A. Ginesi, "DVB-S2 modem algorithms design and performance over typical satellite channels," *Intern. J. of Satellite Communications and Networking*, vol. 22, pp. 281–318, May/June 2004.

[56] B. F. Beidas and R. I. Seshadri, "Analysis and compensation for nonlinear interference of two high-order modulation carriers over satellite link," *IEEE Trans. Commun.*, vol. 58, pp. 1824–1833, June 2010.

[57] B. F. Beidas, "Intermodulation distortion in multicarrier satellite systems: analysis and turbo Volterra equalization," *IEEE Trans. Commun.*, vol. 59, pp. 1580–1590, June 2011.

[58] R. Piazza, B. Shankar, E. Zenteno, D. Ronnow, J. Grotz, F. Zimmer, M. Grasslin, F. Heckmann, and S. Cioni, "Multicarrier digital pre-distortion/equalization techniques for non-linear satellite channels," in *Proc. AIAA Intern. Communications Satellite Systems Conf.*, (Ottawa, Canada), Sept. 2012.

[59] K. P. Liolis and N. S. Alagha, "On 64-APSK constellation design optimization," in *Proc. Intern. Work. on Signal Processing for Space Commun.*, (Rhodes Island, Greece), pp. 1–7, October 2008.

[60] R. M. Gray, *Toeplitz and Circulant Matrices: A review*. Now Publishers Inc, 2006.

[61] F. Charles, *An introduction to the calculus of variations*. Dover, 2010.

Acknowledgement

Probably the *acknowledgement* is my favorite part. The reason is two fold: first, it means that I finished my thesis (yes!). Second, it is the only part of my thesis written with the heart, and not the reason.

So, let me start from Lena. Let be $f : \mathbb{R} \to \mathbb{N}$, such that for a given time t, it is the number of thanks that she deserves, it holds

$$\forall M > 0, M \in \mathbb{N}, \exists t_0 > 0 \text{ s.t. } f(t) > M, \forall t > t_0 \,.$$

(I hope that finally you will understand the limit of a function). I cannot avoid to cite my family and co-family: my parents, my sister, my brothers-in-law with their *lumachina* eating-pizza-all-time, and their crazy frog (impossible write their actual names without a typo, so I will not at all).

Moving to the University: Amina (also well known as RNH_2) does not want to get any acknowledgement, thus I will not. I just let know that she was for me like a co-advisor during my whole PhD. Then, without a particular order, my thanks to Ale, Tommy, Nicolo (this thesis is written in English, thus without *accento*) and the other guys of *Pal 2*.

Cannot forget all the Lund-University guys: Fredrik, Dzevdan, Rohit, Martina, Egle, Taimoor, and everyone else made my stay in Lund wonderful.

Last, but not least my closest friends: Giubbos, *Ingegner duca conte* Pier, Ruben, *Ragionier* Villazzi, and all the other guys of *the company* as well. Also need to cite Maria Paola, lost somewhere in the huge Milan, but without forgetting us in Parma.

And that’s all... just kidding! I did not forget you Giulio! But you know, I’m not leaving now, and this is not a farewell, so let time pass and *put off till tomorrow what you can do today*. Mmm, maybe it was different, but it does not matter.

Andrea

©Piled Higher and Deeper by Jorge Cham, www.phdcomics.com. Permission for publication granted to the author of this thesis, in december 2013.