

Turbo-Coded APSK Modulations Design for Satellite Broadband Communications

Riccardo De Gaudenzi, Albert Guillén i Fàbregas, Alfonso Martinez *

Abstract

This paper investigates the design of power and spectrally efficient coded modulations based on Amplitude Phase Shift Keying (APSK) modulation with application to satellite broadband communications. APSK represents an attractive modulation format for digital transmission over nonlinear satellite channels due to its power and spectral efficiency combined with its inherent robustness against nonlinear distortion. For these reasons APSK has been very recently introduced in the new standard for satellite Digital Video Broadcasting named DVB-S2 [1]. Assuming an ideal rectangular transmission pulse, for which no nonlinear inter-symbol interference is present and perfect pre-compensation of the nonlinearity, we optimize the APSK constellation. In addition to the minimum distance criterion, we introduce a new optimization based on channel capacity; this new method generates an optimum constellation for each spectral efficiency. To achieve power efficiency jointly with low bit error rate (BER) floor we adopt a powerful binary serially concatenated turbo-code coupled with optimal APSK modulations through bit-interleaved coded modulation. We derive tight approximations on the maximum-likelihood decoding error probability, and results are compared with computer simulations. In Ref. [2], the current analysis is complemented with the effects related to satellite nonlinear distortion effects with a band-limited transmission pulse and including demodulator timing,

*R. De Gaudenzi is with European Space Agency (ESA-ESTEC), Noordwijk, The Netherlands, e-mail: Riccardo.De.Gaudenzi@esa.int. A. Guillén i Fàbregas was with European Space Agency (ESA-ESTEC), Noordwijk, The Netherlands. He is currently with Institute for Telecommunications Research University of South Australia, Australia, e-mail: albert.guillen@unisa.edu.au. A. Martinez was with European Space Agency (ESA-ESTEC), Noordwijk, The Netherlands. He is currently with Technische Universiteit Eindhoven, Eindhoven, The Netherlands, e-mail: alfonso.martinez@ieee.org.

amplitude and phase estimation errors. The proposed coded modulation scheme is shown to provide a considerable performance advantage compared to current standards for satellite multimedia and broadcasting systems.

1 Introduction

Satellite communication systems strength lies in their ability to efficiently broadcast digital multimedia information over very large areas [3]. A notable example is the so-called direct-to-home (DTH) digital television broadcasting. Satellite systems also provide a unique way to complement the terrestrial telecommunication infrastructure in scarcely populated regions. The introduction of multi-beam satellite antennas with adaptive coding and modulation (ACM) schemes will allow an important efficiency increase for satellite systems operating at Ku or Ka-band [4]. Those technical enhancements require the exploitation of power- and spectrally-efficient modulation schemes conceived to operate over the satellite nonlinear channel. In this paper we will design high-efficiency 16-ary and 32-ary coded modulation schemes suited for nonlinear satellite channels.

To the authors' knowledge there are few references in the literature dealing with 16-ary constellation optimization over *nonlinear channels*, the typical environment for satellite channels. Previous work showed that 16-QAM does not compare favorably with either Trellis Coded (TC) 16-PSK or uncoded 8-PSK in satellite nonlinear channels [5]. The concept of circular APSK modulation was already proposed thirty years ago by [6], where several non band-limited APSK sets were analyzed by means of uncoded bit error rate bounds; the suitability of APSK for nonlinear channels was also made explicit, but concluded that for single carrier operation over nonlinear channel APSK performs worse than PSK schemes. In the current paper we will revert the conclusion. It should be remarked that [6] mentioned the possibility of modulator pre-compensation but did not provide performance results related to this technique. Foschini [7] optimized QAM constellations using asymptotic uncoded probability of error under average power constraints, deriving optimal 16-ary constellation made of an almost equilateral lattice of triangles. This result is

not applicable to satellite channels. In [8] some comparison between squared QAM and circular APSK over linear channels was performed based on the computation of the error bound parameter, showing some minor potential advantage of APSK. Further work on mutual information for modulations with average and peak power constraints is reported in [9], which proves the advantages of circular APSK constellations under those power constraints. Mutual information performance loss for APSK in peak power limited Gaussian complex channels is reported in [10] and compared to classical QAM modulations; it is shown that under this assumption APSK considerably outperforms QAM in terms of mutual information, the gain particularly remarkable for 16-ary and 64-ary constellations.

Forward Error Correcting codes for our application must combine power efficiency and low BER floor with flexibility and simplicity to allow for high-speed implementation. The existence of practical, simple, and powerful such coding designs for binary modulations has been settled with the advent of turbo codes [11] and the recent re-discovery of Low-Density Parity-Check (LDPC) codes [12]. In parallel, the field of channel coding for non-binary modulations has evolved significantly in the latest years. Starting with Ungerboeck's work on Trellis-Coded Modulation (TCM) [13], the approach had been to consider channel code and modulation as a single entity, to be jointly designed and demodulated/decoded. Schemes have been published in the literature, where turbo codes are successfully merged with TCM [14]. Nevertheless, the elegance and simplicity of Ungerboeck's original approach gets somewhat lost in a series of ad-hoc adaptations; in addition the turbo-code should be jointly designed with a given modulation, a solution impractical for system supporting several constellations. A new pragmatic paradigm has crystallized under the name of Bit-Interleaved Coded Modulation (BICM) [15], where extremely good results are obtained with a standard non-optimized, code. An additional advantage of BICM is its inherent flexibility, as a single mother code can be used for several modulations, an appealing feature for broadband satellite communication systems where a large set of spectral efficiencies is needed.

This paper is organized as follows. Sect. 2 gives the system model under the ideal case of

a rectangular transmission pulse ¹. Sect. 3 gives a formal description of APSK signal sets, describes the maximum mutual information and maximum minimum distance optimization criteria and discusses some of the properties of the optimized constellations. Sect. 4 deals with code design issues, describes the BICM approach, provides some analytical considerations based on approximate maximum-likelihood (ML) decoding error probability bounds, and provides some numerical results. The conclusions are finally drawn in Sect. 5.

2 System Model

The baseband equivalent of the transmitted signal at time t , $s_T(t)$ is given by:

$$s_T(t) = \sqrt{P} \sum_{k=0}^{L-1} x(k) p_T(t - kT_s), \quad (1)$$

where P is the signal power, $x(k)$ is the k -th transmitted symbol, drawn from a complex-valued APSK signal constellation \mathcal{X} , with $|\mathcal{X}| = M$, p_T is the transmission filter impulse response, and T_s is the symbol duration (in seconds), corresponding to one channel use. Without loss of generality, we consider transmission of frames with L symbols. The spectral efficiency R is defined as the number of information bits conveyed at every channel use, and is measured in bits per second per Hertz (bps/Hz).

The signal $s_T(t)$ passes through a high-power amplifier (HPA) operated close to the saturation point. In this region, the HPA shows non-linear characteristics that induce phase and amplitude distortions to the transmitted signal. The amplifier is modeled by a memoryless non-linearity, with an output signal $s_A(t)$ at time t given by:

$$s_A(t) = F(|s_T(t)|) e^{j(\phi(s_T(t)) + \Phi(|s_T(t)|))}, \quad (2)$$

where we have implicitly defined $F(A)$ and $\Phi(A)$ as the AM/AM and AM/PM characteristics of the amplifier for a signal with instantaneous signal amplitude A . The signal amplitude is the instantaneous complex envelope, so that the baseband signal is decomposed as $s_T(t) = |s_T(t)| e^{j\phi(s_T(t))}$.

¹This assumption has been dropped in the paper [2].

In this paper, we assume an (ideal) signal modulating a train of rectangular pulses. These pulses do not create inter-symbol interference when passed through an amplifier operated in the nonlinear region. Under these conditions, the channel reduces to an AWGN, where the modulation symbols are distorted following (2). Let x_A denote the distorted symbol corresponding to $x = |x|e^{j\phi(x)} \in \mathcal{X}$, that is, $x_A = F(|x|)e^{j(\phi(x)+\Phi(|x|))}$. After matched filtering and sampling at time kT_s , the discrete-time received signal at time k , $y(k)$ is then given by,

$$y(k) = \sqrt{E_s}x_A(k) + n(k) \quad k = 0, \dots, L - 1, \quad (3)$$

with E_s the symbol energy, given by $E_s = PT_s$, $x_A(k)$ is the symbol at the k -th time instant, as defined above, and $n(k) \sim \mathcal{N}_{\mathbb{C}}(0, N_0)$ is the corresponding noise sample.

This simplified model suffices to describe the non-linearity up to the nonlinear ISI effect, and allows us to easily design constellation and codes. In the paper [2], the impact of nonlinear ISI has been considered, as well as other realistic demodulation effects such as timing and phase recovery.

3 APSK Constellation Design

In this section we define the generic multiple-ring APSK constellation family. We are interested in proposing new criteria on how to design digital QAM constellations of 16 and 32 points with special emphasis on the behavior for nonlinear channels.

3.1 Constellation Description

M -APSK constellations are composed of n_R concentric rings, each with uniformly spaced PSK points. The signal constellation points x are complex numbers, drawn from a set \mathcal{X} given by:

$$\mathcal{X} = \begin{cases} r_1 e^{j\left(\frac{2\pi}{n_1}i + \theta_1\right)} & i = 0, \dots, n_1 - 1, \quad (\text{ring 1}) \\ r_2 e^{j\left(\frac{2\pi}{n_2}i + \theta_2\right)} & i = 0, \dots, n_2 - 1, \quad (\text{ring 2}) \\ \vdots & \\ r_{n_R} e^{j\left(\frac{2\pi}{n_R}i + \theta_{n_R}\right)} & i = 0, \dots, n_{n_R} - 1, \quad (\text{ring } n_R) \end{cases} \quad (4)$$

where we have defined n_ℓ , r_ℓ and θ_ℓ as the number of points, the radius and the relative phase shift corresponding to the ℓ -th ring respectively. We will nickname such modulations as $n_1 + \dots + n_{n_R}$ -APSK. Fig. 1 depicts the 4+12- and 4+12+16-APSK modulations with quasi-Gray mapping. In particular, for next generation broadband systems [1], [4], the constellation sizes of interest are $|\mathcal{X}| = 16$ and $|\mathcal{X}| = 32$, with $n_R = 2$ and $n_R = 3$ rings respectively. In general, we consider that \mathcal{X} is normalized in energy, i.e., $E[|x|^2] = 1$, which implies that the radii r_ℓ are normalized such that $\sum_{\ell=1}^{n_R} n_\ell r_\ell^2 = 1$. Notice also that the radii r_ℓ are ordered, so that $r_1 < \dots < r_{n_R}$.

Clearly, we can also define the phase shifts and the ring radii in relative terms rather than in absolute terms, as in (4); this removes one dimension in the optimization process, yielding a practical advantage. We let $\phi_\ell = \theta_\ell - \theta_1$ for $\ell = 1, \dots, n_R$ be the phase shift of the ℓ -th ring with respect to the inner ring. We also define $\rho_\ell = r_\ell/r_1$ for $\ell = 1, \dots, n_R$ as the relative radii of the ℓ -th ring with respect to r_1 . In particular, $\phi_1 = 0$ and $\rho_1 = 1$.

3.2 Constellation Optimization in AWGN

We are interested in finding an APSK constellation, defined by the parameters $\boldsymbol{\rho} = (\rho_1, \dots, \rho_{n_R})$ and $\boldsymbol{\phi} = (\phi_1, \dots, \phi_{n_R})$, such that a given cost function $f(\mathcal{X})$ reaches a minimum. The simplest, and probably most natural, cost function is the minimum Euclidean distance between any two points in the constellation. Sect. 3.2.1 shows the results under this criterion. These results are extended in Sect. 3.2.2, where the cost function is replaced by the average mutual information (or channel capacity) of the AWGN channel; it also shown that significant gains may be achieved for low and moderate values of SNR by fine-tuning the constellation.

3.2.1 Minimum Euclidean Distance Maximization

The union bound on the uncoded symbol error probability [16] yields,

$$P_e \leq \frac{1}{M} \sum_{x \in \mathcal{X}} \sum_{\substack{x' \in \mathcal{X} \\ x' \neq x}} Q \left(\sqrt{\frac{E_s |x - x'|^2}{2N_0}} \right), \quad (5)$$

where $Q(x) = 1/\sqrt{2\pi} \int_x^\infty e^{-(t^2/2)} dt$ is the Gaussian tail function. At high signal-to-noise ratio (SNR) Eq. (5) is dominated by the pairwise term at minimum squared Euclidean distance $\delta_{\min}^2 = \min_{x, x' \in \mathcal{X}} |x - x'|^2$. Due to the monotonicity of the Q function, it is clear that maximizing this distance optimizes the error performance estimated with the union bound at high SNR.

The minimum distance of the constellation depends on the number of rings n_R , the number of points in each ring n_1, \dots, n_{n_R} , the radii r_1, \dots, r_{n_R} , and the offset among the rings $\phi_1, \dots, \phi_{n_R}$. The constellation geometry clearly indicates that the distances to consider are between points belonging to the same ring, or between points in adjacent rings. Simple calculations give the following formula:

$$\delta_{\text{ring } i}^2 = 2r_i^2 \left[1 - \cos\left(\frac{2\pi}{n_i}\right) \right] \quad (6)$$

for the distance between points in ring i -th, with radius r_i and n_i points. For the adjacent rings the calculation is only slightly more complicated, and gives the following:

$$\delta_{\text{rings } i, i+1}^2 = r_i^2 + r_{i+1}^2 - 2r_i r_{i+1} \cos \theta \quad (7)$$

where θ is the minimum relative offset between any pair of points of rings i and $i + 1$ respectively. As the phase of point l_i in ring i is given by $\phi_i + 2\pi l_i/n_i$, we easily obtain:

$$\theta = \min_{l_i, l_{i+1}} \left| \left(\phi_i - \phi_{i+1} \right) + 2\pi \left(\frac{l_i}{n_i} - \frac{l_{i+1}}{n_{i+1}} \right) \right|. \quad (8)$$

The minimum distance of the constellation is given by taking the minimum of all these inter-ring and intra-ring values:

$$\delta_{\min}^2 = \min_{\substack{i=1, \dots, n_R \\ j=1, \dots, n_R-1}} \{ \delta_{\text{ring } i}^2, \delta_{\text{rings } j, j+1}^2 \}. \quad (9)$$

For the sake of space limitations, we concentrate on 16-ary constellations. Thanks to symmetry considerations, it is clear that the best offset between rings happens when $\phi_2 = \pi/n_2$. Fig. 2 shows the minimum distance for several candidates: 4+12-, 6+10-, 5+11- and 1+5+10-APSK. It may be observed that the highest minimum distance is achieved for approximately $\rho_2 = 2.0$, except for 4+12-APSK, where $\rho_2 = 2.7$. The results for $\phi = 0$ are also plotted, and show that

the corresponding minimum distance is smaller. We will see later in Sect. 3.2.2 how this effect translates into error rate performance.

3.2.2 Mutual Information Maximization

The average mutual information (assuming equiprobable symbols) for a given signal set \mathcal{X} provides the maximum transmission rate (in bits/channel use) at which error-free transmission is possible with such signal set, and is given by (e.g. [15]),

$$f(\mathcal{X}) = C = \log_2 M - \mathbb{E}_{x,n} \left\{ \log_2 \left[\sum_{x' \in \mathcal{X}} \exp \left(-\frac{1}{N_0} |\sqrt{E_s}(x - x') + n|^2 - |n|^2 \right) \right] \right\}. \quad (10)$$

Interestingly, for a given signal-to-noise ratio, or equivalently, for a given spectral efficiency R , an optimum constellation can be obtained, a procedure we apply in the following to 16- and 32-ary constellations.

In general, closed-form optimization of this expression is a daunting task, so we resort to numerical techniques. Expression (10) can be easily evaluated by using the Gauss-Hermite quadrature rules, making numerical evaluation very simple. Note, however, that it is possible to calculate a closed-form expression for the asymptotic case $E_s/N_0 \rightarrow +\infty$. First, note that the expectation in Eq. (10) can be rewritten as:

$$\lambda(\mathcal{X}) \triangleq \mathbb{E}_{x,n} \left\{ \log_2 \left[\sum_{x' \in \mathcal{X}} \exp \left(-\frac{1}{N_0} \left(E_s |x - x'|^2 + 2\text{Re}(\sqrt{E_s}(x - x')n) \right) \right) \right] \right\}. \quad (11)$$

Using the dominated convergence theorem [17], the influence of the noise term vanishes asymptotically, since the limit can be pushed inside the expectation. Furthermore, the only remaining terms in the summation over $x' \in \mathcal{X}$ are $x' = x$ and those closest in Euclidean distance $\delta_{\min}^2 = \min_{x' \in \mathcal{X}} |x - x'|^2$, of which there are $n_{\min}(x)$. Therefore the expectation becomes:

$$\lambda(\mathcal{X}) \simeq \mathbb{E}_x \left\{ \log_2 \left[1 + n_{\min}(x) \exp \left(-\frac{1}{N_0} E_s \delta_{\min}^2 \right) \right] \right\} \quad (12)$$

Noting that the exponential takes very small values, the approximation $\log_2(1+x) \simeq x \log_2 e$ for $|x| \ll 1$ holds, thus by simplifying further the expectation we obtain:

$$\lambda(\mathcal{X}) \simeq \mathbb{E}_x \left\{ n_{\min}(x) \exp\left(-\frac{E_s}{N_0} \delta_{\min}^2\right) \log_2 e \right\} \simeq \alpha \exp\left(-\frac{E_s}{N_0} \delta_{\min}^2\right). \quad (13)$$

where α is a constant that does not depend on the constellation minimum distance δ_{\min} nor on SNR. Then the capacity at large SNR becomes:

$$f(\mathcal{X}) = \log_2 M - \alpha \exp\left(-\frac{E_s}{N_0} \delta_{\min}^2\right). \quad (14)$$

It appears then clear that the procedure corresponds to the maximization of the minimum Euclidean distance, as in Sect. 3.2.1.

Fig. 3 shows the numerical evaluation of Eq. (10) for a given range of values of ρ_2 and $\phi = \phi_2 - \phi_1$ for the 4+12-APSK constellation at $E_s/N_0 = 12$ dB. Surprisingly, there is no sensible dependence on ϕ . Therefore, the two-dimensional optimization can be done by simply finding the ρ_2 that maximizes channel capacity. This result was found to hold true also for the other constellations and hence, in the following, capacity optimization results do not account for ϕ . Fig. 4 shows the union bound on the symbol error probability (5) for several 16-APSK modulations, and for the optimum value of ρ_2 at $R = 3$ bps/Hz (found with the mutual information analysis). Continuous lines indicate $\phi = 0$ while dotted lines refer to the maximum value of the relative phase shift, i. e. $\phi = \pi/n_2$, showing no dependence on ϕ at high SNR. This absence of dependency is justified by the fact that the optimum constellation separates the rings by a distance larger than the number of points in the ring itself, so that the relative phase ϕ has no significant impact in the distance spectrum of the constellation.

For 16-APSK it is also interesting to investigate the capacity dependency on n_1 and n_2 . Fig. 5(a) depicts the capacity curves for several configurations of optimized 16-APSK constellations and compared with classical 16-QAM and 16-PSK signal sets. As we can observe, capacity curves are very close to each other, showing a slight advantage of 6+10-APSK over the rest. In particular, note that there is a small gain, of about 0.1 dB, in using the optimized constellation for every R , rather

than the calculated with the minimum distance (or high SNR). However, as discussed in [2], 6+10-APSK and 1+5+10-APSK show other disadvantages compared to 4+12-APSK for phase recovery and nonlinear channel behavior.

Similarly, Fig. 5(b) reports capacity of optimized 4+12+16-APSK (with the corresponding optimal values of ρ_2 and ρ_3) compared to 32-QAM and 32-PSK. We observe slight capacity gain of 32-APSK over PSK and QAM constellations. Other 32-APSK constellations with different distribution of points in the three rings did not provide significantly better results.

Finally Table 5 provides the optimized 16- and 32-APSK parameters for various coding rates, giving an optimum constellation for each given spectral efficiency.

3.3 Constellation Optimization for Nonlinear Channels

3.3.1 Peak-to-Envelope Considerations

For nonlinear transmission over an amplifier, 4+12-APSK is preferable to 6+10-APSK because the presence of more points in the outer ring allows to maximize the HPA DC power conversion efficiency. It is better to reduce the number of inner points, as they are transmitted at a lower power, which corresponds a lower DC efficiency. It is known that the HPA power conversion efficiency is monotonic with the input power drive up to its saturation point. Fig. 5 shows the distribution of the transmitted signal envelope for 16-QAM, 4+12-APSK, 6+10-APSK, 5+11-APSK, and 16-PSK. In this case the shaping filter is a square-root raised cosine (SRRC) with a roll-off factor $\alpha = 0.35$ as for the DVB-S2 standard [1]. As we observe, the 4+12-APSK envelope is more concentrated around the outer ring amplitude than 16-QAM and 6+10-PSK, being remarkably close to the 16-PSK case. This shows that the selected constellation represents a good trade-off between 16-QAM and 16-PSK, with error performance close to 16-QAM, and resilience to nonlinearity close to 16-PSK. Therefore, 4+12+APSK is preferable to the rest of 16-ary modulations considered. Similar advantages have been observed for 32-APSK compared to 32-QAM.

3.3.2 Static Distortion Compensation

The simplest approach for counteracting the HPA nonlinear characteristic for the APSK signal, as already introduced in Sect. 2, is to modify the complex-valued constellation points at the modulator side. Thanks to the multiple-ring nature of the APSK constellation, pre-compensation is easily done by a simple modification of the parameters ρ_ℓ , and ϕ_ℓ . The objective is to exploit the known AM/AM and AM/PM HPA characteristics in order to obtain a good replica of the desired signal constellation geometry after the HPA, as if it had not suffered any distortion. This can be simply obtained by artificially increasing the relative radii ρ_ℓ and modifying the relative phases ϕ_ℓ at the modulator side. This approach neglects nonlinear ISI effects at the matched filter output which are not present under the current assumption of rectangular symbols; ISI issues has been discussed in [2].

In the 16-ary APSK case the new constellation points x' follow (4), with new radii r'_1, r'_2 , such that $F(r'_1) = r_1$, and $F(r'_2) = r_2$. Concerning the phase, it is possible to pre-correct for the relative phase offset introduced by the HPA between inner and outer ring by simply changing the relative phase shift by $\phi'_2 = \phi_2 + \Delta\phi$, with $\Delta\phi = \phi(r'_2) - \phi(r'_1)$. These operations can be readily implemented in the digital modulator by simply modifying the reference constellation parameters ρ', ϕ' , with *no hardware complexity impact* or out-of-band emission increase at the linear modulator output. On the other side, this allows to shift all the compensation effort into the modulator side allowing the use of an optimal demodulator/decoder for AWGN channels even when the amplifier is close to saturation. The pre-compensated signal expression at the modulator output is then,

$$s_T^{\text{pre}} = \sqrt{P} \sum_{k=0}^{L-1} x'(k) p_T(t - kT_s) \quad (15)$$

where now $x'(k) \in \mathcal{X}'$ being the pre-distorted symbols with r'_ℓ and ϕ'_ℓ for $\ell = 1, \dots, n_R$.

4 Forward Error Correction Code Design and Performance

In this section we describe the coupling of turbo-codes and the APSK signal constellations through BICM and we discuss some of the properties of this approach ². As already mentioned in Sect. 1, such approach is a good candidate for flexible constellation format transmission. The main drivers for the selection of the FEC code have been flexibility, i.e. use a single mother code, independently of the modulation and code rates; complexity, i.e. have a code as compact and simple as possible; and good performance, i.e. approach Shannon's capacity bound as much as possible.

We consider throughout a coded modulation scheme for which the transmitted symbols $\mathbf{x} = (x_0, \dots, x_{L-1})$ are obtained as follows: 1) The information bits sequence $\mathbf{a} = (a_0, \dots, a_{K-1})$ is encoded with a binary code $\mathcal{C} \in \mathbb{F}_2^N$ of rate $r = K/N$; 2) The encoded sequence $\mathbf{c} = (c_0, \dots, c_{N-1}) \in \mathcal{C}$ is bit-interleaved, with an index permutation $\boldsymbol{\pi} = (\pi_0, \dots, \pi_{N-1})$; 3) The bit-interleaved sequence \mathbf{c}_π is mapped to a sequence of modulation symbols \mathbf{x} with a labeling rule $\mu : \mathbb{F}_2^M \rightarrow \mathcal{X}$, such that $\mu(a_1, \dots, a_M) = x$. In addition to the description of the code, we also propose the use of some new heuristics to tune the final design of the BICM codes.

4.1 Code Design

It was suggested in [15] that the binary code \mathcal{C} can be optimized for a binary channel (such as BPSK or QPSK with AWGN). We substantiate this claim with some further insights on the effect of the code minimum distance in the error performance. The Bhattacharyya union bound (BUB) on the frame error probability P_e for a BICM modulation assuming that no iterations are performed at the demapper side is given by [15]:

$$P_e \leq \sum_d A(d)B(E_s/N_0)^d, \quad (16)$$

²The optimization method based on the mutual information proposed in Sect. 3.2.2 can be easily extended to the case of the BICM mutual information [15] with almost identical results assuming Gray mapping. However, we use the proposed method in order to keep the discussion general and not dependent on the selected coding scheme.

where $A(d)$ is the number of codewords at a Hamming distance d , d_{\min} is the minimum Hamming distance, with $B(E_s/N_0)$ denoting the Bhattacharyya factor, which is given by:

$$B(E_s/N_0) = \frac{1}{M \log_2 M} \sum_{i=1}^{\log_2 M} \sum_{b=0}^1 \sum_{x \in \mathcal{X}_{i=b}} \mathbb{E}_n \left\{ \sqrt{\frac{\sum_{z \in \mathcal{X}_{i=b}} \exp(-\frac{1}{N_0} |x - z + n|^2)}{\sum_{z \in \mathcal{X}_{i=b}} \exp(-\frac{1}{N_0} |x - z + n|^2)}} \right\}. \quad (17)$$

Eq. (17) can be evaluated very efficiently using the Gauss-Hermite quadrature rules. For sufficiently large E_s/N_0 the BUB in Eq. (5) is dominated by the term at minimum distance, i. e., the error floor

$$P_e \simeq A_{d_{\min}} B(E_s/N_0)^{d_{\min}}. \quad (18)$$

From this equation we can derive an easy lower bound on the d_0 on the minimum distance of \mathcal{C} for a given target error rate, modulation, and number of codewords at d_{\min} :

$$d_{\min} \geq \lceil d_0 \rceil, \quad \text{where } d_0 = \frac{\log P_e - \log A_{d_{\min}}}{\log B}, \quad (19)$$

where $\lceil x \rceil$ denotes the smallest integer greater or equal to x . Notice that the target error rate is fixed to be the error floor under ML decoding³. The lowest error probability floor is achieved by a code \mathcal{C} with $A_{d_{\min}} = 1$. Fig. 7 shows the lower bound d_0 with $A_{d_{\min}} = 1$, as a function of E_s/N_0 for target $P_e = 10^{-4}, 10^{-7}$, QPSK, 16-QAM, 16-APSK and 32-APSK modulations and Gray labelling. In order to ease the comparison, a normalized SNR is used, defined as:

$$\left. \frac{E_s}{N_0} \right|^{\text{norm}} = \frac{E_s}{N_0} \frac{1}{2^R - 1} \quad (20)$$

where R is the spectral efficiency, and the normalization is thus to the channel capacity. The code rate has been taken $r = 3/4$ for all cases. Note that a capacity-achieving pair modulation-code would work at a normalized $\left. E_s/N_0 \right|^{\text{norm}} = 1$, or 0 dB.

A remarkable conclusion is that BICM with Gray mapping preserves the properties of \mathcal{C} regardless of the modulation used, since we observe that the requirements for non-binary modulations

³Although this does not necessarily hold under iterative decoding, it does still provide a useful guideline into the performance.

are strikingly similar to those for binary modulations (in the error-floor region). In order to work at about 3 dB from capacity, that is, a normalized $E_s/N_0|_{\text{norm}} = 3$ dB, the needed d_0 is about 5 and 10 for a frame error rate of 10^{-4} and 10^{-7} respectively.

We consider that \mathcal{C} is a serial concatenation of convolutional codes (SCCC) [18], with outer code \mathcal{C}^O of length L_O and rate r_O and inner code \mathcal{C}^I of length L_I and rate r_I . Obviously, $L_I = N$ and $r_O r_I = r$. The resulting spectral efficiency is $R = r \log_2 M$. It provides two key advantages with respect to parallel turbo codes: lower error floor, possibly achieving the bit error rate requirements ($\text{BER} \leq 10^{-10}$) without any external code; and simpler constituent codes simpler than in turbo codes or in classical concatenated codes. In addition, with an SCCC the outer code is fully integrated into the decoding process, which includes iterations between decoding stages for the inner and outer codes. This avoids the need to use an additional external code, such as a Reed-Solomon (and its associated interleaver). In some sense, the outer code is already included in the SCCC code, thus saving one extra encoding/decoding step, and one memory level, therefore reducing the required complexity.

The best choice in terms of low error floor forces the lowest possible rate for the outer encoder, as this maximizes the interleaver gain, which increases exponentially with the outer code free distance [18]. We should then set the outer code rate equal to the total code rate, and the inner code rate to 1. Also, it turns out that the best choice for the inner encoder is the two-state differential encoder also known as “accumulator”. It meets the requirements of simplicity, it is “almost” systematic, in the sense that the dependency among the bits in its output sequence is very mild, and moreover, it is recursive as imposed by the design rules of SCCCs for the inner encoder. Last but not least, this choice leads to a very simple inner SISO, a highly desirable feature for a design working at high data rates.

In practice the maximum block length to be used shall be selected accounting for the maximum allowed end-to-end latency and decoder complexity. One recent finding [19] allows to split an arbitrary block interleaver in an arbitrary number of smaller non-overlapping interleavers. This

allows to greatly reduce the decoder complexity when parallel SISO units are used to achieve high-speed decoding as memory requirements does not increase with the degree of parallelism.

As an outer code, we have selected the standard binary 16-state convolutional code, rate 3/4 [20]. Its free distance is 4, large enough so that interleaving gain can be achieved, and the minimum distance of the concatenated code grows towards infinity with the blocklength [21]. Furthermore, and if required, the code may be punctured to higher rates [22], with no loss in the code distance. Further numerical results are presented in Sect. 4.3.

4.2 Demodulation

Decoding of BICM consists of a concatenation of two steps, namely maximum-a-posteriori (MAP) soft-input soft-output (SISO) demapper (symbol-to-bit likelihood computer), and a MAP SISO decoder of \mathcal{C} . These two steps exchange extrinsic information messages $m_{\mu \rightarrow \mathcal{C}}$ (from the demapper to the SISO decoder of \mathcal{C}), and $m_{\mathcal{C} \rightarrow \mu}$ (from the SISO decoder of \mathcal{C} to the demapper) through the iterations. Extrinsic information messages m (or metrics) can be in the form of likelihood probabilities, log-likelihood ratios or some combination or approximation of them. When either \mathcal{C}^O or \mathcal{C}^I , or both, are convolutional codes, MAP SISO decoding is efficiently computed by the BCJR algorithm [23]. For example, the extrinsic log-likelihood ratio corresponding to $m_{\mu \rightarrow \mathcal{C}}$ for the i -th coded bit of the k -th symbol and l -th iteration is given by,

$$\Lambda \left(c_{k,i}^{(l)} \right) = \log \frac{\sum_{x \in \mathcal{X}_{i=0}} p(y_k|x) \prod_{j \neq i} P_{\mathcal{C} \rightarrow \mu}^{(l-1)}(c_{k,j})}{\sum_{x \in \mathcal{X}_{i=1}} p(y_k|x) \prod_{j \neq i} P_{\mathcal{C} \rightarrow \mu}^{(l-1)}(c_{k,j})} \quad (21)$$

where $p(y_k|x) \propto \exp \left(-\frac{1}{N_0} |y_k - \sqrt{E_s} x|^2 \right)$, $P_{\mathcal{C} \rightarrow \mu}^{(l)}(c)$ denotes the extrinsic probability message corresponding to $m_{\mathcal{C} \rightarrow \mu}$ on the coded bit c at the l -th iteration, and $\mathcal{X}_{i=b} = \{x \in \mathcal{X} | \mu_i^{-1}(x) = b\}$, where $\mu_i^{-1}(x) = b$ denotes that the i -th position of binary label x is equal to b .

There is a marginal information loss in considering no iterations at the demodulator side when Gray mapping is used for transmitting high rates [15], i.e., $P_{\mathcal{C} \rightarrow \mu}^{(l-1)}(c_{k,j}) = 0.5$ implies almost no loss in spectral efficiency using Gray mapping. When demapper iterations are allowed, Gray

mapping is known not to gain through the iterations [24]. Moreover, when other mapping rules are used, scheduling the operations for such decoder (a SCC with BICM) can be a very complicated task and has been solved only for $N \rightarrow \infty$ (see e. g. [25] for recent results on the subject). For all these aforementioned reasons, we will assume Gray mapping and that information flows from demodulator to decoder only, with no feed-back.

4.3 Performance Analysis

Density evolution (or approximations such as EXIT charts [24]) of such turbo-coded BICM is a very complicated task due to the concatenation of three elements exchanging extrinsic information messages through the iterations. Such techniques lead in general to 3-dimensional surfaces which are difficult to deal with in practical decoding algorithms for finite length codes [25]. We will therefore resort to a mixture of computer simulations and bounds on Maximum Likelihood (ML) decoding error probability. Regarding convergence, simulations can accurately estimate the values of E_b/N_0 at which the decoding algorithm does not converge, as will be shown shortly.

We denote the binary-input channel between the modulator and demodulator as the equivalent binary-input BICM channel. It has been recently shown [26] that such channel can be very well approximated as AWGN⁴ with SNR $\gamma = -\log B(E_s/N_0)$. Therefore, standard bounds for binary-input channels can be successfully applied here. In particular, the standard union bound (UB) yields

$$P_e \lesssim \sum_d A_d Q\left(\sqrt{-2d \log B(E_s/N_0)}\right). \quad (22)$$

At high SNR, (16) and (22) are dominated by the pairwise terms corresponding to the few codewords with low Hamming distance (error floor). When turbo-like codes are used, union bounding techniques are known not to provide good estimates of the error probability, and one typically

⁴Notice that the Gaussian approximation (GA) is common practice in density evolution techniques [24].

resorts to improved bounds such as the tangential-sphere bound (TSB) [27],

$$P_F \lesssim \int_{-\infty}^{+\infty} \frac{dz_1}{\sqrt{2\pi\sigma^2}} e^{-z_1^2/2\sigma^2} \left\{ 1 - \bar{\Gamma} \left(\frac{N-1}{2}, \frac{r_{z_1}}{2\sigma^2} \right) + \sum_{d: \delta/2 < \alpha_\delta} A_d \bar{\Gamma} \left(\frac{N-2}{2}, \frac{r_{z_1}^2 - \beta_\delta(z_1)^2}{2\sigma^2} \right) \left[Q \left(\frac{\beta_\delta(z_1)}{\sigma} \right) - Q \left(\frac{r_{z_1}}{\sigma} \right) \right] \right\}, \quad (23)$$

where $\bar{\Gamma}(a, x) = \frac{1}{\Gamma(a)} \int_0^x t^{a-1} e^{-t} dt$ is the normalized incomplete gamma function and $\Gamma(x) = \int_0^{+\infty} t^{x-1} e^{-t} dt$ is the gamma function, $\sigma^2 = (-2 \log B(E_s/N_0))^{-1}$, $r_{z_1} = r(1 - z_1/4R)$, $\beta_\delta(z_1) = \frac{r_{z_1}}{\sqrt{1 - \delta^2/4R^2}} \frac{\delta}{2r}$, $\alpha_\delta = r\sqrt{1 - \delta^2/4R^2}$, $R^2 = N$, $\delta = 2\sqrt{d}$ and r , the cone radius, is the solution of $\sum_{d: \delta/2 < \alpha_\delta} A_d \int_0^{\theta_k} \sin^{N-3} \phi d\phi = \frac{\sqrt{\pi} \Gamma(\frac{N-2}{2})}{\Gamma(\frac{N-1}{2})}$ with $\theta_k = \cos^{-1} \left(\frac{\delta}{2r} \frac{1}{\sqrt{1 - \delta^2/4R^2}} \right)$.

Fig. 8 shows the bit-error probability bounds using the Gaussian approximation, Eq. (22) and simulations for BPSK, 4+12-APSK and 4+12+16-APSK with the pseudo-Gray labelings of Fig. 1. In particular, in Fig. 8(a) we use a repeat and accumulate (RA) code [28] with $r = 1/4$ and $K = 512$ information bits per frame. In this case, the weight enumerator can be computed in closed form [28]. We observe that as expected BICM preserves the properties of the underlying binary code \mathcal{C} , since both waterfall and floor occur at almost the same probability values. We also observe that the approximations in Eqs. (22) and (23) are very accurate and yield much better error probability estimates than the standard Bhattacharyya bound. Same conclusions apply to Fig. 8(b) where we use a SCCC with the optimal 16 states $r = 3/4$ convolutional code as outer code and inner accumulator, with interleaver size $N = 5000$. Storage limitations prevent us from showing the curves for larger N . However, we observe that already with $N = 5000$ we have very low error floors. In particular, in the DVB-S2 application, the interleaver size is set to 16200 or 64000, which implies that almost-error-free transmission is possible with such code.

Finally, Fig. 9, shows the simulated BER performance for the same SCCC with the optimal 16 states $r = 3/4$ convolutional code as outer code and inner accumulator, with interleaver size $N = 16200^5$ with 4+12-APSK, 16-QAM, 4+12+16-APSK and 32-QAM. For the sake of comparison,

⁵The selected FEC block size ensures that the FEC floor is well below the required BER of 10^{-10} for satellite

we also plot the BER for a 4+12-APSK with a 16 states TCM, typical of satellite systems current standard [3]. As we observe, the SCCC codes yield a substantial performance improvement with respect to TCM. In the TCM case, one usually concatenates a Reed Solomon code operating at an input BER of 10^{-4} , which usually diminishes the spectral efficiency and increases the receiver complexity. Notice as well that 32-APSK achieves a better performance than 32-QAM, giving a further justification to the use of modulations in the APSK family instead of the classical QAM.

5 Summary and Conclusions

Extensive analysis and simulations for turbo-coded APSK modulations, with particular emphasis on its applicability to satellite broadband communications have been presented in this paper. In particular, we have investigated APSK constellation optimization under mutual information and minimum Euclidean distance criteria, under the simplified assumption of rectangularly shaped transmission pulses. We have shown that the degrees of freedom in the design of an APSK modulation can be exploited thanks to the mutual information maximization, and this has been applied to the design of 16- and 32-ary constellations. This technique has been shown to extend the standard minimum Euclidean distance maximization, yielding a small but significant improvement.

The pragmatic approach of BICM allows for a good coupling between such optimized APSK modulations with powerful binary turbo-codes, due to its inherent flexibility for multiple-rate transmission. Some new heuristics have been used to further justify the design of a single mother code to be used for all rates. A theoretical explanation of the the fact that the error floor typical of turbo codes remains at a constant distance from capacity has been presented. We have presented some new ML decoding error probability bounds for BICM APSK, and we have compared them with simulations findings. Numerical results based on simulation of bit-error rate probability for high rate transmission with turbo-coded APSK have been presented, showing large advantage of the presented scheme over standard TCM.

broadcasting systems.

References

- [1] *ETSI Digital Video Broadcasting (DVB); Second generation framing structure, channel coding and modulation systems for Broadcasting, Interactive Services, News Gathering and other broadband satellite applications*, European Telecommunications Standards Institute Std. EN 302 307, 2003.
- [2] R. De Gaudenzi, A. Guillén i Fàbregas, and A. Martinez, “Performance analysis of turbo-coded APSK modulations over nonlinear satellite channels,” *IEEE Trans. Wireless Commun.*, accepted for publication, 2005.
- [3] M. Cominetti and A. Morello, “Digital video broadcasting over satellite (DVB-S): a system for broadcasting and contribution applications,” *Intl. J. Satell. Commun.*, no. 18, pp. 393–410, 2000.
- [4] R. De Gaudenzi and R. Rinaldo, “Adaptive coding and modulation for next generation broadband multimedia systems,” in *Proc. of the 20th AIAA Satellite Communication Systems Conference*, Montreal, May 2002.
- [5] E. Biglieri, “High-level modulation and coding for nonlinear satellite channels,” *IEEE Trans. Commun.*, vol. COM-32, no. 5, May 1984.
- [6] C. M. Thomas, M. Y. Weidner, and S. H. Durrani, “Digital amplitude-phase keying with m-ary alphabets,” *IEEE Trans. Commun.*, vol. COM-22, no. 2, pp. 168–180, February 1974.
- [7] G. Foschini, R. Gitlin, and S. Weinstein, “Optimization of two-dimensional signal constellations in the presence of Gaussian noise,” *IEEE Trans. Commun.*, vol. 22, no. 1, pp. 28–38, January 1974.
- [8] G. Einarsson, “Signal design for the amplitude-limited Gaussian channel by error bound optimization,” *IEEE Trans. Commun.*, vol. 27, no. 1, pp. 152–158, January 1979.
- [9] S. Shamai and I. Bar-David, “The capacity of average and peak-power-limited quadrature Gaussian channels,” *IEEE Trans. Inform. Theory*, vol. 41, pp. 1060–1071, July 1995.
- [10] R. Müller, U. Wachsmann, and J. Huber, “Multilevel coding for peak power limited complex Gaussian channels,” in *Proc. IEEE International Symposium on Information Theory (ISIT '97)*, June-July 1997, p. 103.
- [11] C. Berrou, A. Glavieux, and P. Thitimajshima, “Near Shannon-limit error correcting coding and decoding: Turbo codes,” in *Proc. IEEE Int. Conference on Communications (ICC'93)*, Geneva, Switzerland, May 1993.
- [12] R. Gallager, “Low-density parity-check codes,” Ph.D. dissertation, Massachusetts Institute of Technology, 1963.
- [13] G. Ungerboeck, “Channel coding with multilevel phase signals,” *IEEE Trans. Inform. Theory*, vol. IT-28, January 1982.

- [14] P. Robertson and T. Worz, "Bandwidth-efficient turbo trellis-coded modulation using punctured component codes," *IEEE J. Select. Areas Commun.*, vol. 16, no. 2, pp. 206–218, February 1998.
- [15] G. Caire, G. Taricco, and E. Biglieri, "Bit-interleaved coded modulation," *IEEE Trans. Inform. Theory*, vol. 44, pp. 927–947, May 1998.
- [16] S. Benedetto, E. Biglieri, and V. Castellani, *Digital Transmission Theory*. Prentice Hall, 1987.
- [17] A. Dembo and O. Zeitouni, *Large Deviations Techniques and Applications*, 2nd ed., ser. Applications of Mathematics. Springer Verlag, April 1998, no. 38.
- [18] S. Benedetto, D. Divsalar, G. Montorsi, and F. Pollara, "Serial concatenated of interleaved codes: Performance analysis, design and iterative decoding," *IEEE Trans. Inform. Theory*, vol. 44, no. 3, pp. 909–926, May 1998.
- [19] A. Tarable, G. Montorsi, and S. Benedetto, "Mapping interleaver laws to parallel turbo decoder architectures," in *Proc. 3th Intl. Symposium on Turbo Codes and Applications*, Brest, France, September 2003.
- [20] J.-J. Chang, D.-J. Hwang, and M.-C. Lin, "Some extended results on the search for good convolutional codes," *IEEE Trans. Inform. Theory*, vol. 43, no. 5, pp. 1682–1697, September 1997.
- [21] N. Kahale and R. Urbanke, "On the minimum distance of parallel and serially concatenated codes," 1997, accepted *IEEE Trans. on Information Theory*.
- [22] A. Martinez and B. Ponticelli, "Efficient coding schemes for high order modulations in satellite broadcasting applications," in *Proc. 7th Intl. ESA Workshop on Digital Signal Processing Techniques for Space Commun.*, Sesimbra, Portugal, October 2001.
- [23] L. Bahl, J. Cocke, F. Jelinek, and J. Raviv, "Optimal decoding of linear codes for minimizing the symbol error rate," *IEEE Trans. Inform. Theory*, vol. 20, pp. 284–287, March 1974.
- [24] S. ten Brink, "Designing iterative decoding schemes with the extrinsic information transfer chart," *AEÜ Int. J. Electron. Commun.*, December 2000.
- [25] F. Brannstrom, L. K. Rasmussen, and A. Grant, "Optimal scheduling for multiple serially concatenated systems," in *Proc. 3th Intl. Symposium on Turbo Codes and Applications*, Brest, France, September 2003.
- [26] A. Guillén i Fàbregas, A. Martinez, and G. Caire, "Error probability of bit-interleaved coded modulation using the Gaussian approximation," in *Proc. of the 38th Annual Conference on Information Sciences and Systems (CISS) 2004*, Princeton, US, March 2004.
- [27] I. Sason and S. Shamai, "Improved upper bounds on the ML decoding error probability of parallel and serial concatenated turbo codes via their ensemble distance spectrum," *IEEE Trans. Inform. Theory*, vol. 46, no. 1, pp. 24–47, January 2000.

- [28] D. Divsalar, H. Jin, and R. J. McEliece, “Coding theorems for ‘turbo-like’ codes,” in *Proc 36th Allerton Conf. on Communication, Control and Computing*, Allerton, Illinois, US, September 1998, pp. 210–219.

Modulation order	Coding rate r	Spectral eff. (b/s/Hz)	ρ_1^{opt}	ρ_2^{opt}
4+12-APSK	2/3	2.67	3.15	N/A
4+12-APSK	3/4	3.00	2.85	N/A
4+12-APSK	4/5	3.20	2.75	N/A
4+12-APSK	5/6	3.33	2.70	N/A
4+12-APSK	8/9	3.56	2.60	N/A
4+12-APSK	9/10	3.60	2.57	N/A
4+12+16-APSK	3/4	3.75	2.84	5.27
4+12+16-APSK	4/5	4.00	2.72	4.87
4+12+16-APSK	5/6	4.17	2.64	4.64
4+12+16-APSK	8/9	4.44	2.54	4.33
4+12+16-APSK	9/10	4.50	2.53	4.30

Table 1: Capacity optimized constellation parameters for 16-ary and 32-ary APSK

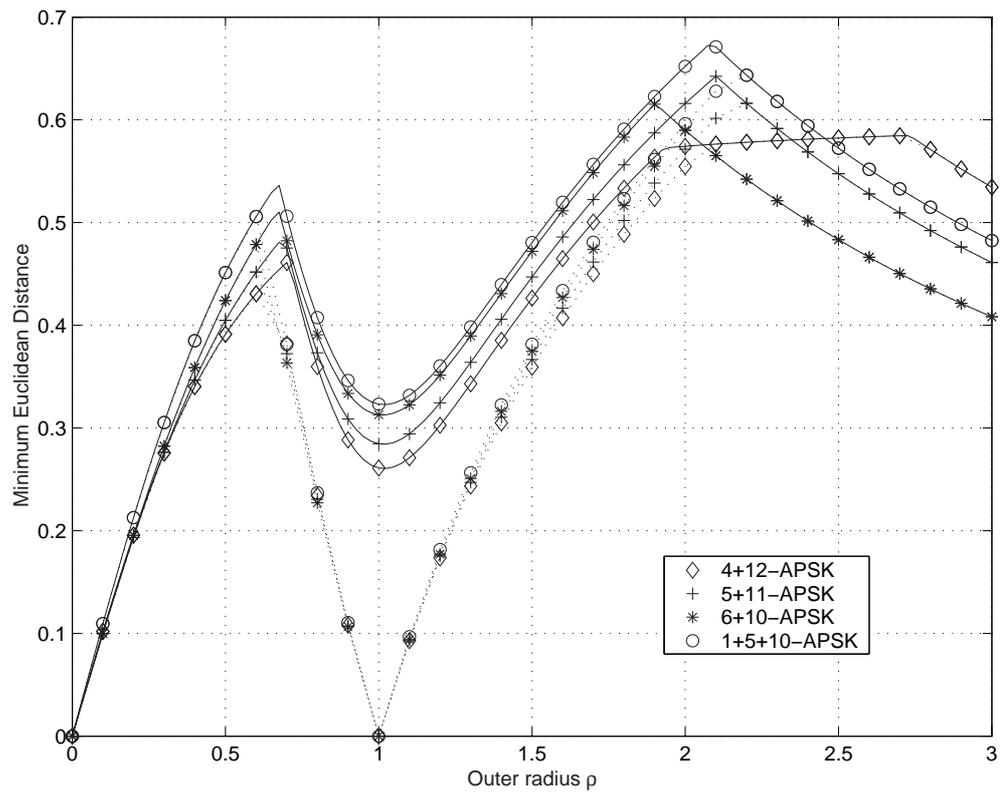


Figure 2: Minimum Euclidean distances for several 16-ary signal constellations. Solid lines correspond to $\phi = \pi/n_2$; dotted lines to $\phi = 0$.

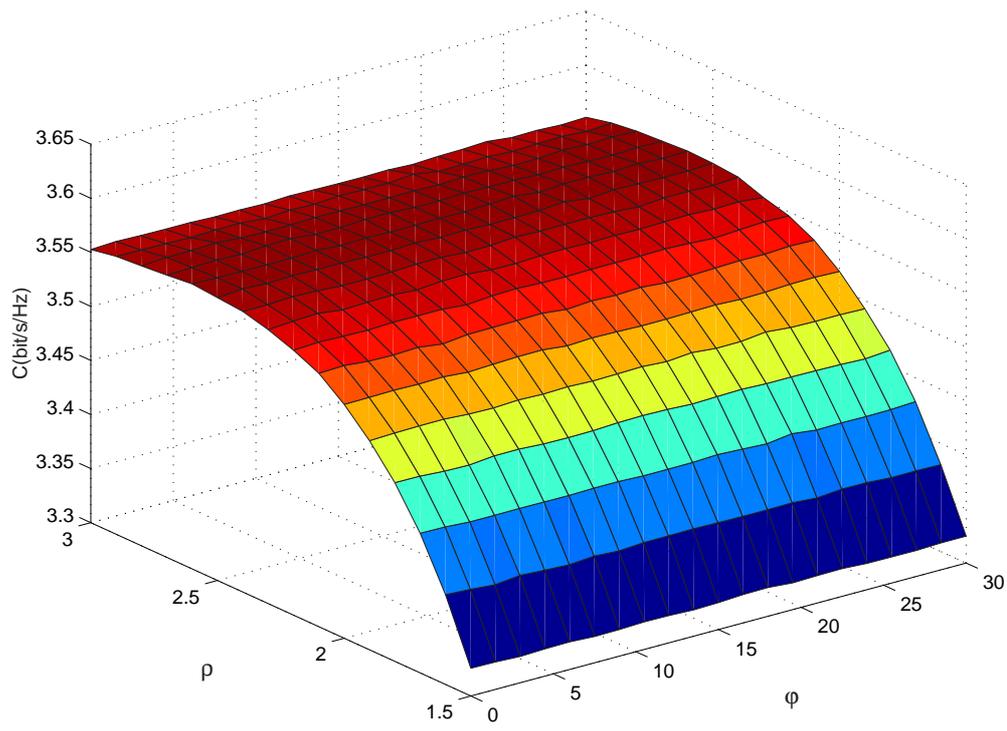


Figure 3: Capacity surface for the 16-APSK ($n_1 = 4, n_2 = 12$), with $E_s/N_0 = 12$ dB.

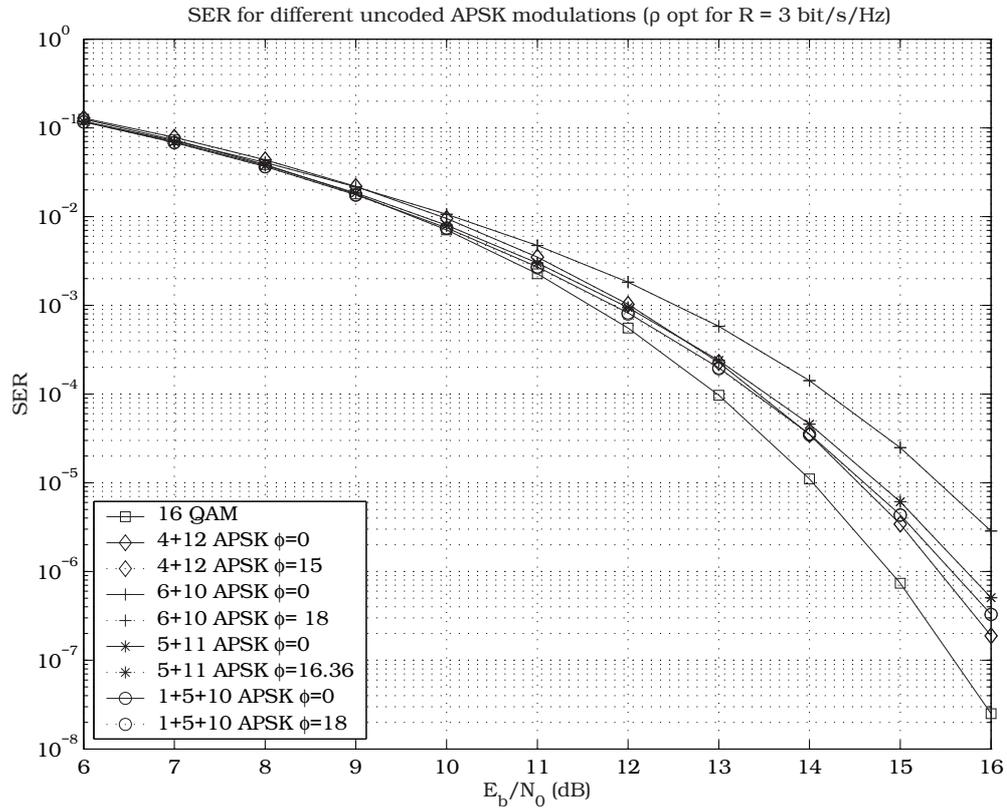
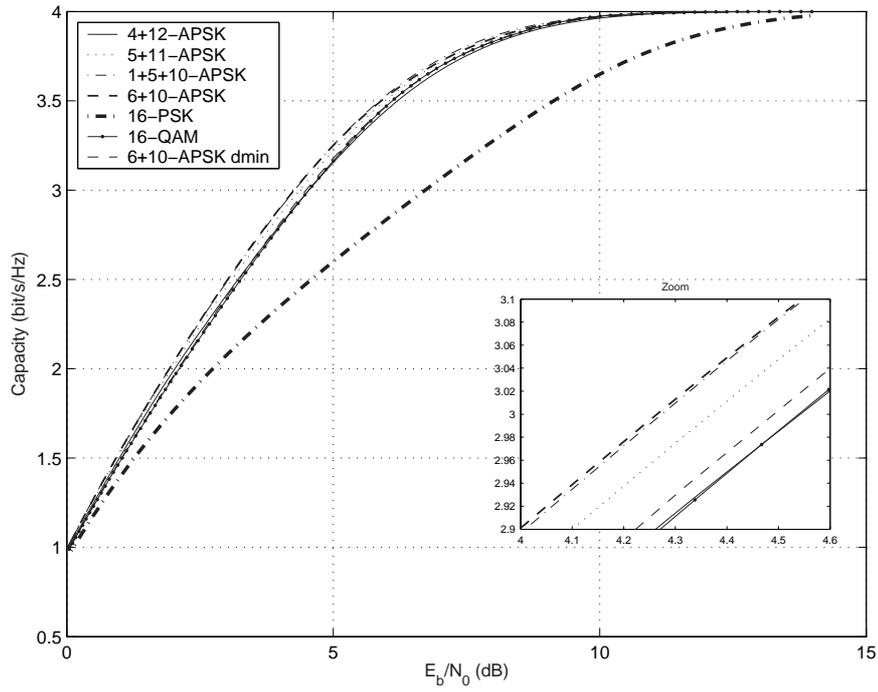
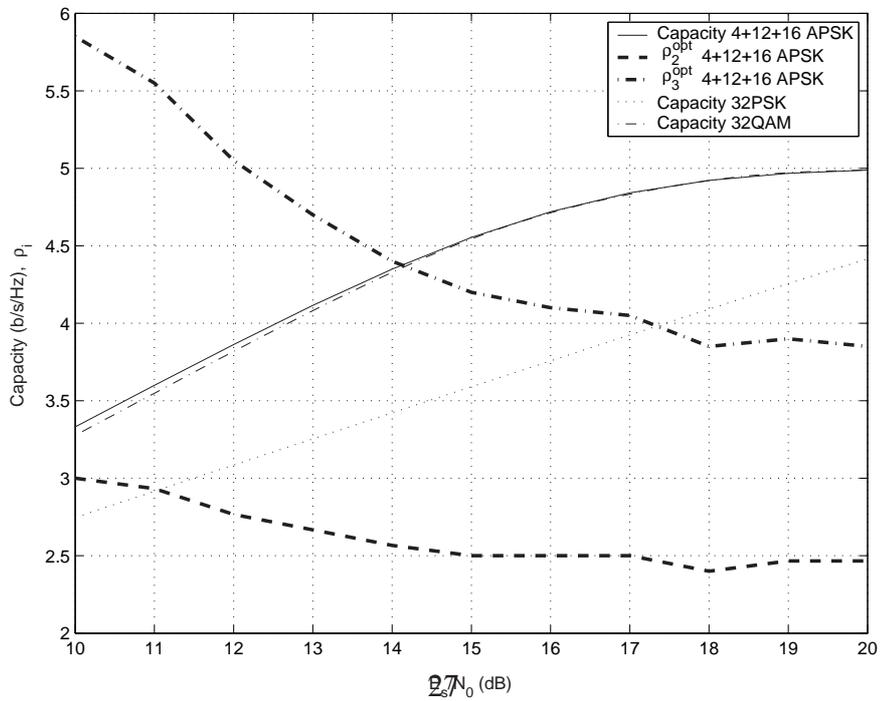


Figure 4: Union bound on the uncoded symbol error probability for several APSK modulations. Note that the continuous line and the dashed line are indistinguishable because they are superimposed.



(a) 16-ary constellations (zoom around 3 bps/Hz).



(b) 32-ary constellations.

Figure 5: Capacity and ρ^{opt} for the optimized APSK signal constellations versus QAM and PSK.

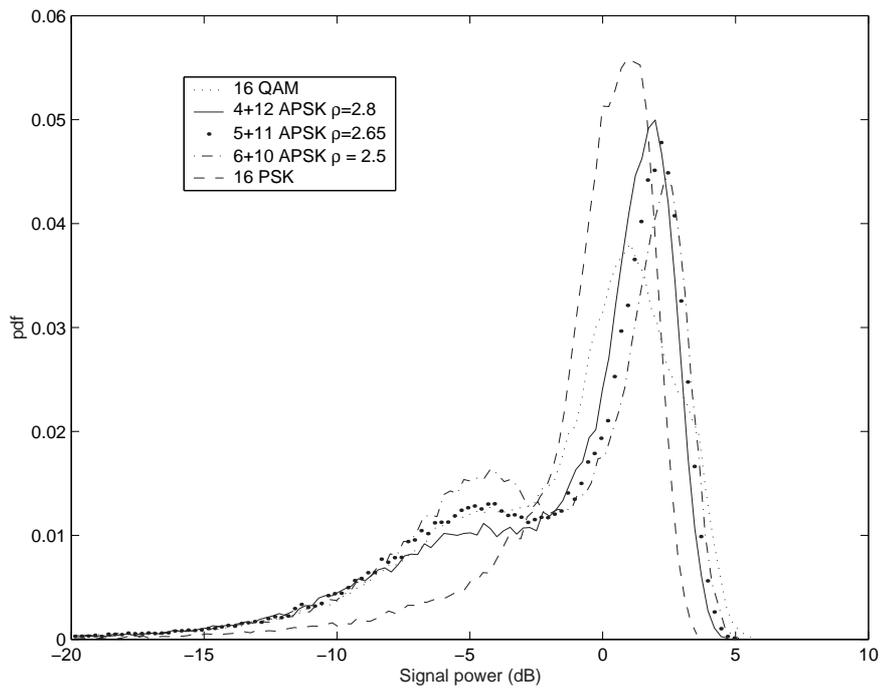


Figure 6: Simulated histogram of the transmitted signal envelope power for 16-ary constellations.

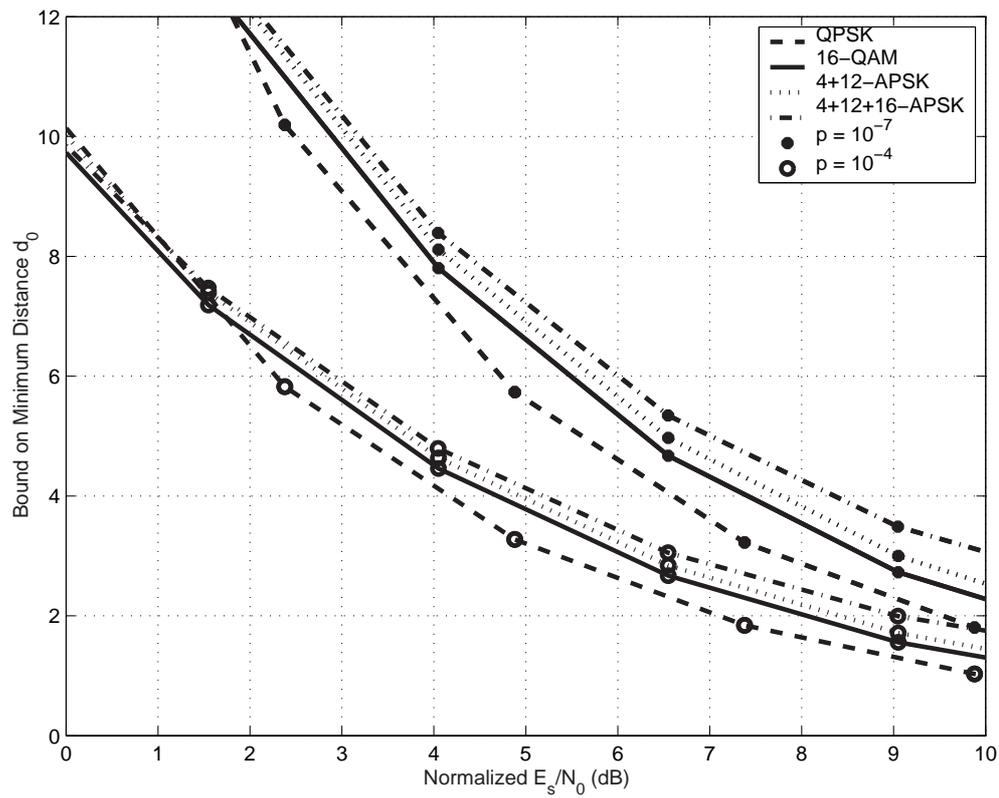
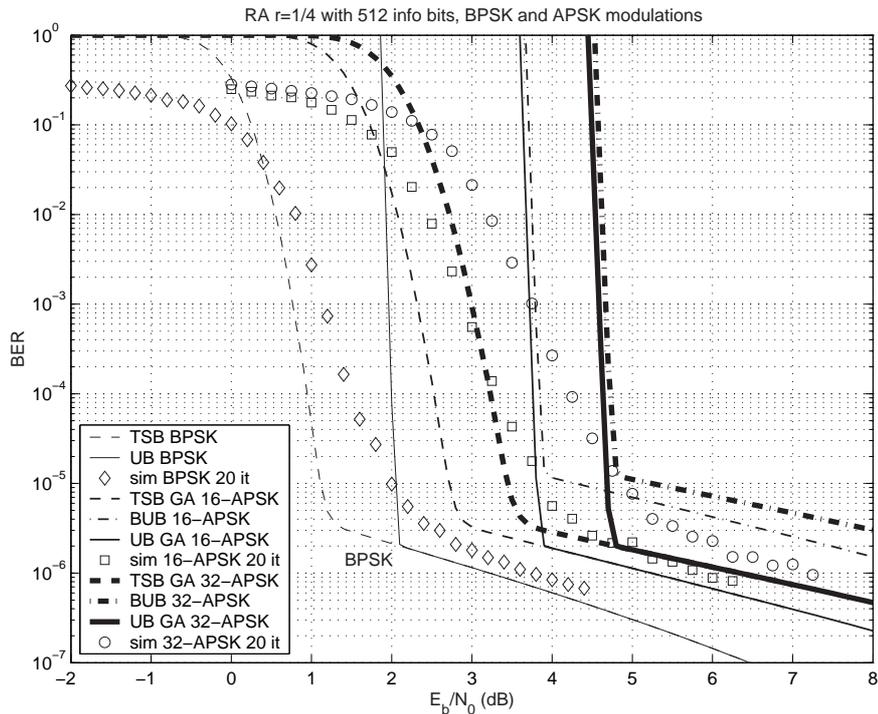
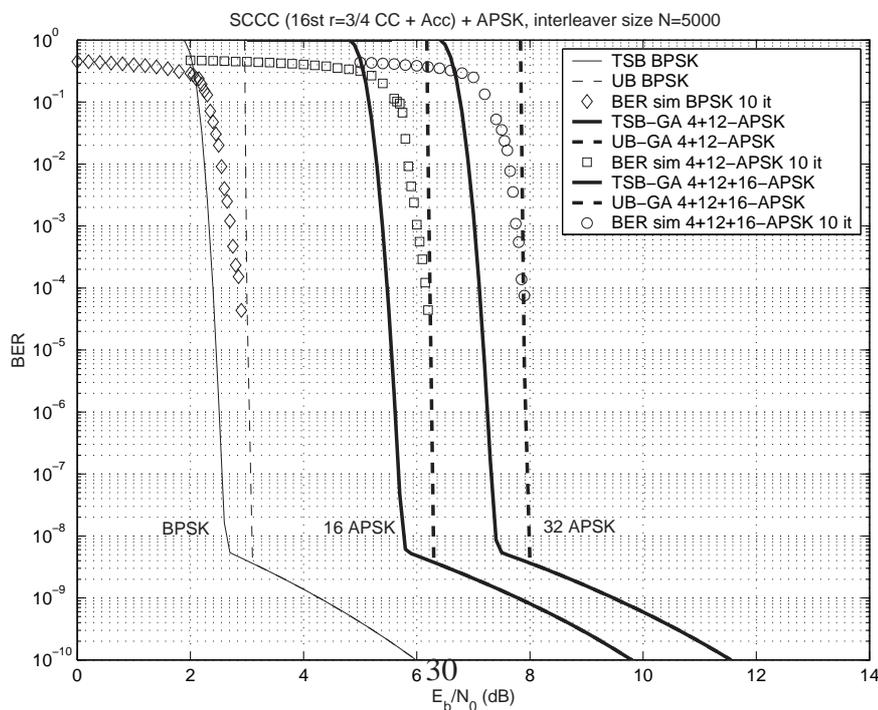


Figure 7: Lower bound d_0 vs. normalized E_s/N_0 for target $P_e = 10^{-4}, 10^{-7}$, QPSK and 16- and 32-ary modulations and Gray labelling.



(a) RA code with $r = 1/4$ and $K = 512$ information bits per frame



(b) SCCC with rate $3/4$ 16 states convolutional code as outer code and inner accumulator with $N = 5000$

Figure 8: Bit-error probability bounds and simulations for BPSK and 16-APSK ($n_1 = 4$ and $n_2 = 12$) and 32-APSK ($n_1 = 4$, $n_2 = 12$ and $n_3 = 16$) with pseudo-Gray labelling.

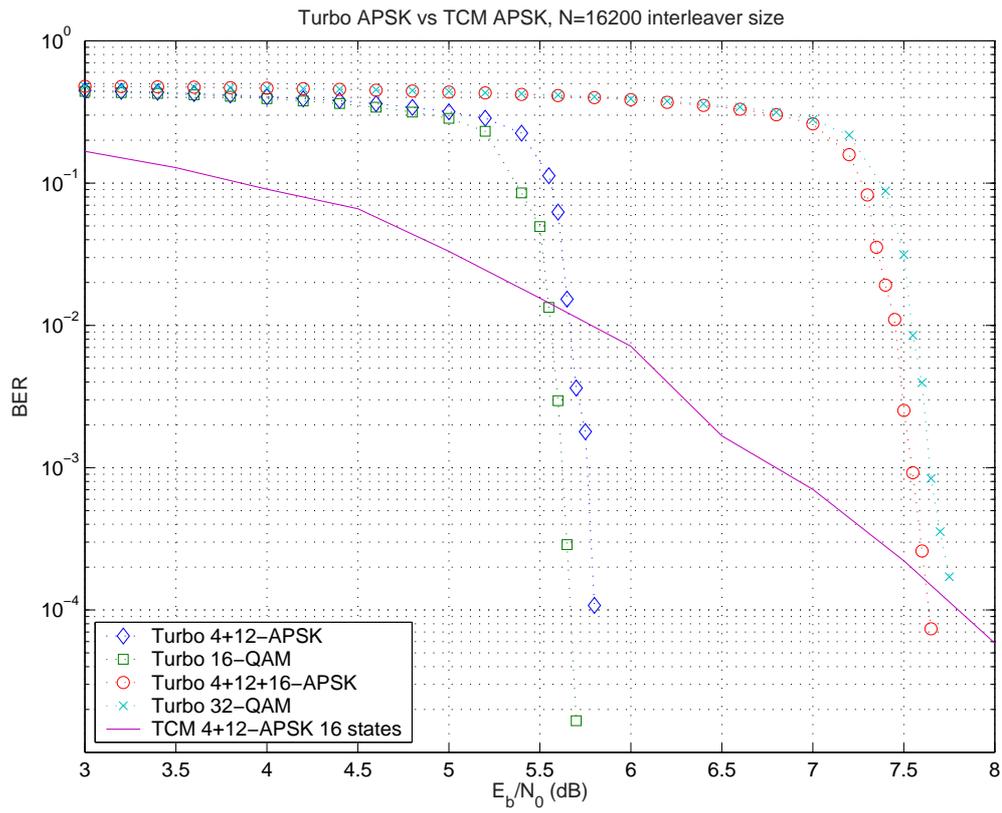


Figure 9: Turbo-coded APSK with $N = 16200$ vs 4+12-TCM.