APSK CODED MODULATION SCHEMES FOR NONLINEAR SATELLITE CHANNELS WITH HIGH POWER AND SPECTRAL EFFICIENCY

Riccardo De Gaudenzi,* Alfonso Martinez Vicente, Beatrice Ponticelli
European Space Agency ESTEC, Keplerlaan 1, 2200 AG Noordwijk, The Netherlands

Albert Guillén i Fàbregas†
Institut EURECOM, 2229 Route des Cretes, 06904 Sophia Antipolis, France

A new class of 16-ary Amplitude Phase Shift Keying (APSK) coded modulations deemed double-ring PSK modulations best suited for (satellite) nonlinear channels is proposed. Constellation parameters optimization has been based on geometric and information-theoretic considerations. Furthermore, pre- and post-compensation techniques to reduce the nonlinearity impact have been examined. Digital timing clock and carrier phase have been derived and analyzed for a Turbo coded version of the same new modulation scheme. Finally, the performance of state-of the art Turbo coded modulation for this new 16-ary digital modulation has been investigated and compared to the known TCM schemes. It is shown that for the same coding scheme, double-ring APSK modulation outperforms classical 16-QAM and 16-PSK over a typical satellite nonlinear channel due to its intrinsic robustness against the High Power Amplifier (HPA) nonlinear characteristics. The new modulation is shown to be power- and spectrally-efficient, with interesting applications to satellite communications.

INTRODUCTION

Despite the growing fiber capabilities to carry ultra-high speed digital information, satellite communications undoubtedly still represent a big success thanks to its capability to broadcast digital multi-media information[1] over very large regions or to complement lack of terrestrial infrastructure. This is the case for satellite news gathering systems linking mobile TV stations to the central production facility[1] or the emerging multi-media satellite systems under development.[2] In the search for power and spectral efficient modulation able to suit next generation communication satellite requirements two main directions can be envisioned; development of highly efficient coding schemes based on the turbo principle supporting variable coding rates, and efficient coupling with multi-dimensional modulations suited to the satellite nonlinear channel.

It is well known that for satellite channels the constellation 16-QAM (Quadrature Amplitude Modulation), although providing twice the spectral efficiency of the widely used QPSK (Quadrature Phase Shift Keying), suffers greatly from satellite amplifier nonlinearity effects. Trellis coded (TC) 16-QAM has been adopted for high speed satellite links (e.g. TV contributions see[1]) at the cost of higher operating link signal-to-noise ratios (E_b/N_0 increase of typically 4-4.5 dB) and a larger operational amplifier output back-off (OBO) (typically 3 to 5 dB). The current demand for bandwidth efficient high-speed satellite communication links calls for the development of more effective alternatives to 16-QAM. In the following a very power and bandwidth efficient 16-ary coded modulation for nonlinear (satellite) channels dubbed 16-APSK is introduced and its performance analyzed.

CONSTELLATION OPTIMIZATION

In this section, 16-ary double ring PSK modulations are described, and its parameter optimization performed. It has been observed that 16-QAM suffers from severe amplitude and phase distortions due to the satellite Traveling Wave Tube Amplifier (TWTA) nonlinear characteristics. The effect is even more evident for general M-QAM constellations, with $M > 16$. As already noticed in,[6] when the High Power Amplifier (HPA) is driven close to the saturation point, it tends to compress the 16-QAM squared constellation onto a "double ring" constellation, i.e., a constellation composed by a 4-PSK lying on the inner ring and (non uniform) 12-PSK lying on an outer ring. This is due to the saturation of the amplifier output, which compresses

*Directorate of Technical and Operational Support, Communication Systems Section, e-mail: Riccardo.De.Gaudenzi@esa.int

†Research Fellow, Mobile Communications Department

the external points of the 16-QAM constellation towards a constant amplitude ring. The post-compensation in the trellis decoder metric calculation proposed in[6] although noticeably improving the BER performance does not allow to avoid the losses due to the distorted constellation shape.

The idea developed here is to jointly optimize the transmitted signal constellation, taking into account the nonlinear channel nature together with the potential demodulator performance improvement through modulation predistortion.

As depicted in Fig. 1, we define a $N_1 + N_2$-PSK modulation as a double ring PSK modulation where N_1, R_1, N_2, R_2 are the number of points and corresponding radii of the inner and outer rings respectively, $\rho = R_2/R_1$, φ as the relative phase shift between both PSK constellations, and $N_1 R_1^2 + N_2 R_2^2 = 1$ is the unit energy normalization condition. The above double-ring complex signal envelope

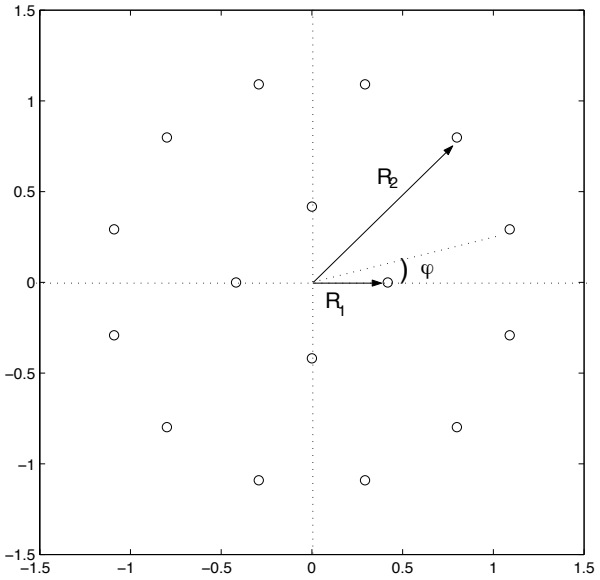

Fig. 1 Parametric description of double ring 16-APK constellation 4+12-PSK: $N_1 = 4$, $N_2 = 12$, $\rho = 2.7$ **and** $\varphi = 15$.

expression in time domain is:

$$\tilde{s}_{TX}(t) = \sqrt{P} \sum_{k=-\infty}^{\infty} R(\psi_k) \exp[j\psi_k]\ g_T(t - kT_s) \quad (1)$$

where P is the signal power, $R(\psi_k) \in \{R_1, R_2\}$ is the constellation signal amplitude alphabet, $\psi_k \in \{\Psi_1^1, \cdots, \Psi_{N_1}^1, \Psi_1^2, \cdots, \Psi_{N_2}^2\}$ with $N_1 + N_2 = M$ is the constellation phase alphabet, $g_T(t)$ is the transmission filter impulse response and $T_s = T_b' \log_2 M$ is the channel symbol duration while $T_b' = T_b r$ is the coded bits duration, being r the coding rate and T_b the bit interval of the information source. In particular the following relation between the phase alphabet and the ring amplitude holds:

$$R(\psi_k) = \begin{cases} R_1 & \text{for } \psi_k \in \{\Psi_1^1, \cdots, \Psi_{N_1}^1\} \\ R_2 & \text{for } \psi_k \in \{\Psi_1^2, \cdots, \Psi_{N_2}^2\} \end{cases} \quad (2)$$

Constellation parameters optimization can be done by following two different approaches. The first and most intuitive way of looking at the problem, implies the computation of all the possible distances between all different pairs of points, determine the *minimum distance* d_{min} as function of ρ and φ, and determine for which ρ and φ, d_{min} is maximal. This approach, however, does not take into account the whole constellation geometric properties. The performance, in a general case, does depend on the set of all distances between pairs of constellation points, not only the minimal, and therefore, a more general approach should be envisaged. A well-established result from Information Theory is the channel capacity, which, for a given modulation, provides the maximum achievable rate for which transmission is possible with vanishing error probability.

In particular, the capacity of two dimensional APSK constellations over a band-limited AWGN channel with no inter-symbol interference will be optimized as a function of N_1, N_2, ρ and φ,

$$(N_1, N_2, \rho, \varphi)_{opt} = \arg \max_{N_1, N_2, \rho, \varphi} C(N_1, N_2, \rho, \varphi). \quad (3)$$

In the case where we use a specific modulation constellation, namely a set of M complex numbers $a_k, k = 1, \ldots, M$, the channel capacity becomes,[4]

$$\begin{aligned} C &= \log_2 M - \\ &- \frac{1}{M} \sum_{i=0}^{M-1} E\left[\log_2 \sum_{j=0}^{M-1} \exp\left(- \frac{|a_i + n - a_j|^2 - |n|^2}{N_0} \right) \right] \end{aligned} \quad (4)$$

where, a_i and a_j are the complex signal representations of two constellation points, with average signal E_s, and n is a complex noise random variable with variance $N_0/2$. The channel symbol SNR in terms of the capacity C and the average energy per bit to noise spectral density E_b/N_0 is $E_s/N_0 = C(E_s/N_0)E_b/N_0$.

Parameter optimization is done by computing the channel capacity for a given $N_1 + N_2$-PSK modulation as function of ρ and φ, which for the 4+12-PSK case, gives the capacity surface shown in Fig. 2. In this case ,over AWGN linear channel, the capacity is optimized by $\rho = 2.7$. It was also shown that there is little sensitivity to the relative

inner/outer ring phase offset φ. Another interesting candidate is the so called 6+10-PSK ($N_1 = 6$, and $N_2 = 10$), for which the optimum is $\rho = 2.21$.[5]

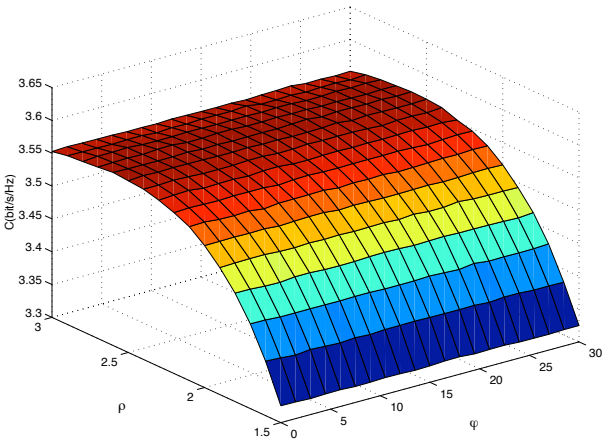

Fig. 2 Capacity surface for 4+12-PSK as function of ρ and φ.

The curves of channel capacity for the optimized 16-APSK modulations are shown in Fig. 3, where it may be observed that the capacity bounds for 16-QAM, 4+12-PSK and 6+10-PSK are very close for the AWGN linear channel. Taking into account the previous discussion, we can expect that 4+12-PSK or 6+10-PSK modulation formats are likely to outperform 16-QAM over the nonlinear channel. For the sake of completeness, optimum[7] hexagonal constellations (HEX-A, HEX-B) slightly outperforming 16-QAM are reported, even though they are clearly not suited for the nonlinear channel.

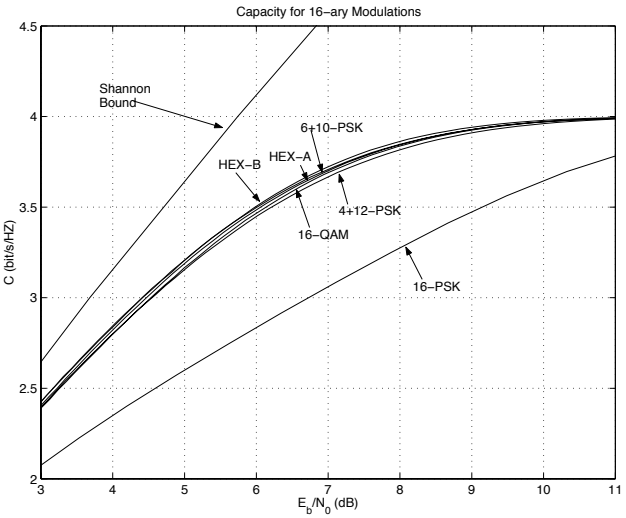

Fig. 3 Capacity for the two optimized 16-APK signal constellations versus 16-QAM and 16-PSK.

Despite the result of Fig. 3 showing a slight capacity advantage for 6+10-PSK with respect to 4+12-PSK, considering the nonlinear channel application, the latter constellation has been preferred as the presence of more points in the outer ring will allow to maximize the HPA DC efficiency. This is because the inner points will be transmitted at a lower power, to which corresponds a lower DC efficiency. Fig. 4 shows the distribution of the transmitted signal envelope for four constellations, namely, 16-QAM, 4+12-PSK, 6+10-PSK and 16-PSK. The shaping filter is a square-root raised cosine with a roll-off factor $\alpha = 0.35$.

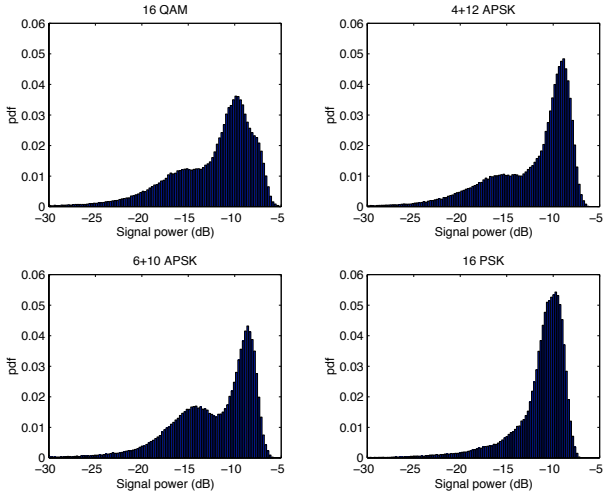

Fig. 4 Simulated histogram of the transmitted signal envelope power: a) 16-QAM, b) 4+12-PSK, $\rho = 2.7$, $\varphi = 0$, c) 6+10-PSK, $\rho = 2.21$, $\varphi = 0$, d) 16-PSK.

It can be noticed that the 4+12-PSK envelope is more concentrated around the (common) outer ring amplitude than 16-QAM and 6+10-PSK, being remarkably close to the 16-PSK case. This shows that the selected constellation, 4+12-PSK, 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.

SATELLITE CHANNEL

Satellite Channel Model

For simplicity in the following we will assume a simple transmission chain composed of a digital modulator, square-root raised cosine (SRRC) band-limiting (with roll-off factor 0.35) and a high power amplifier characterized by a typical AM/AM and AM/PM Ka-band TWTA characteristic. This scheme is also representative of a satellite bent-pipe transponder for which the uplink noise is negligible compared to the downlink. Due to the tight signal band-limiting the impact of the satellite output analog filter is negligible. The satellite channel is then simply represented by a memory-less nonlinearity. For the numerical

examples in the following we will consider the AM/AM and AM/PM characteristics shown in Fig. 5.

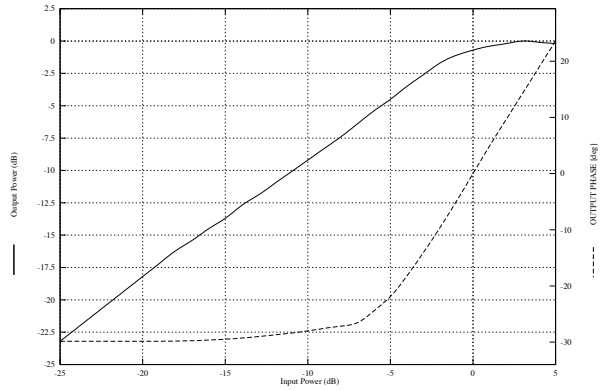

Fig. 5 I-O characteristics of the reference Ka-band TWT Amplifier.

We introduce here the parameter $[E_b/N_0]_{\rm sat}$ defined as the ratio between the transmitted energy per bit when the amplifier is driven at saturation by a continuous wave (CW) carrier and the noise power spectral density.[3] It is easy to find that the following general relation holds when a digital signal is passed through the HPA driven at a given input back-off (IBO):

$$[E_b/N_0]_{\rm sat} = [E_b/N_0]_{\rm inp}({\rm IBO})({\rm dB}) + {\rm OBO}({\rm IBO})({\rm dB}) \quad (5)$$

It means that the effective demodulator input $[E_b/N_0]_{\rm inp}$ is reduced by the output back-off (OBO) with respect to the one potentially available for a system operating with a single constant envelope signal operating at HPA saturation. At the same time the demodulator performance is degraded with respect to an ideal linear AWGN channel by an amount D(IBO) (dB) (dependent on the HPA distortion and hence on the IBO/OBO) so that the effective demodulator input E_b/N_0 named $[E_b/N_0]_{\rm eff}$ is given by:

$$[E_b/N_0]_{\rm sat} = [E_b/N_0]_{\rm eff}({\rm dB}) + {\rm OBO}({\rm IBO})({\rm dB}) + {\rm D}({\rm IBO})({\rm dB}) \quad (6)$$

This equation is very useful to optimize the performance of a nonlinear digital transmission system as it allows to optimize the HPA IBO (and/or OBO) that minimizes $[E_b/N_0]_{\rm sat}$. The optimal operating point represents the best trade-off between the increasing power loss (OBO) related to the increasing IBO and the reduction of the distortion (D) due to the improved linearity experienced by an increasing IBO.

Satellite channel distortion compensation

Several techniques to mitigate the effect of satellite channel nonlinearity are possible. Pre-compensation techniques attempt to counteract the HPA distortion through signal (constellation) pre-distortion. Post-compensation techniques allow to recover distortion losses at the demodulator side through nonlinear equalization and/or *ad-hoc* decoder metric computation. For comprehensive critical review of the possible solutions one can refer to.[5]

In the following we concentrate on a simple technique just requiring the APSK generated signal parameter modification. The key observation leading to this proposal is that a reduction in the distortion can be obtained by changing the complex-valued constellation, instead, of the transmitted signal. In fact, the non-linearity has two major effects:

– The creation of inter-symbol interference at the receiver, due to the non-matched filter receiver. This issue is to be tackled mainly with an equalizer at the receiver, or at the transmitter by means of shaping pulse optimization.

– The distortion of the very constellation points, which are mapped to a different point. If the distortion were linear, this would amount to a shift in the constellation, recoverable with phase and amplitude estimation. For a non-linear case, the relative positions of the constellation points change. This can be reduced by pre-compensation at the transmitter, or post-compensation at the receiver.

It is quite easy to see that thanks to the double-ring APSK constellation properties, one can easily pre-compensate for the main HPA constellation geometry distortion effects by simply modifying the generated signal ρ and φ parameters. This avoids the use of nonlinear pre-compensation devices after the modulator acting on the transmitted signal. Taking into account the AM/AM and AM/PM HPA characteristics it is easy to see that a good approximation to the original signal constellation geometry can be re-obtained by artificially increasing the ρ factor in the modulator. This assumes one sample per symbol after the matched filter and neglects inter-symbol interference effects. In this way the optimal value of ρ, for instance 2.7 for 4+12-PSK, is re-established at the demodulator side. The outer ring phase φ is also de-rotated to counteract the HPA AM/PM distortion effects. If required, the pre-compensation parameters can be modified to track possible HPA characteristic aging effects. The major advantage compared to the post-compensation technique discussed in[6] is that for a broadcasting system the compensation only affects the central (ground) modulator and does not impact the distributed demodulator parameters.

This type of APSK pre-compensation can be readily implemented in the digital modulator by simply modifying the reference constellation parameters $\rho^{'}, \varphi^{'}$ *with no hardware complexity impact* or out-of band emission increase at the modulator output. On the other side this allows to shift all the compensation effort onto the modulator side allowing the use of an optimal demodulator/decoder for AWGN channels even when the amplifier is operated very close to saturation.

TRELLIS-CODED PERFORMANCE

Overall system description

The end-to-end TCM system under consideration is described in Fig. 6. The binary information data bits b_k at rate R_b enter a serial-to-parallel device (S/P) generating three parallel streams at rate $R_b/3$. The rate $r = 3/4$ trellis encoder generates four parallel binary symbol streams at rate $R_s = R_b^{'}/(\log_2 M)$, with $R_b^{'} = R_b/r$, that are mapped through an Ungerboeck set partitioning mapping to the 16-ary constellation generator. The I-Q multilevel digital pulse stream is then passed to the two baseband SRRC filters and I-Q modulated at RF. In case of the nonlinear channel the passband real signal then drives the HPA whose model has been described earlier. Additive White Gaussian Noise (AWGN) representative of the downlink satellite channel is then added. The passband demodulator input signal is I-Q demodulated to baseband using standard analog or digital techniques. The I-Q streams are then asynchronously sampled at $N_s = T_s/T_{clock}$ samples/symbol. In practical implementations the number of samples/symbol N_s depends on the anti-aliasing filter contained in the I-Q demodulator front-end. These samples, assumed for simplicity to be not quantized, are then passed to the receiver SRRC matched filters. The SRRC filter outputs are then subsampled at two samples per symbol (on-time and early) by a digital interpolator[14] driven by the timing error detector unit described in the synchronization section. The output symbol stream is then split into on-time and early samples through a simple S/P converter. Both streams are used to derive the timing error information. The on-time one sample per symbol stream then enters the digital vector tracker described in the synchronization section performing amplitude and phase estimation. We assume here that there is no need for a frequency estimator to help the phase estimator. If necessary, the QAM frequency estimators[12] are applicable to our case. In case the nonlinear equalizer[16] is activated it is inserted prior to the TCM decoder, that represents the last demodulator block. As discussed, tentative symbol decisions are generated to help the digital vector tracker. The ambiguity resolution block takes care of resolving possible phase estimator ambiguities as discussed in.[6]

Code Selection

The overall BER performance of the demodulator for the 16-state rate 3/4 trellis code reported in[10] was successfully compared to the theoretical bound derived by means of the so-called distance-spectrum approach introduced by Rouanne and Costello in.[9] The performance of the 16-QAM demodulator in the linear channel, including amplitude, timing and phase estimation subsystem degradations, is very close to the analytical upper bound in AWGN. For the 16-state code the coding gain of coded 16-QAM with respect to uncoded 8-PSK at $P_e = 10^{-5}$ is about 4.4 dB. Concerning the operating point for the TCM, for quasi-error free (QEF) performance (e.g. BER on the order of 10^{-10}) the inner decoder BER should be on the order of $2{\cdot}10^{-41}$ assuming that an RS (188, 204) code with proper interleaving is adopted as outer code. The selected 16-state optimal trellis code for 16-QAM provides a performance very close to that of the 64-state binary pragmatic trellis code selected for the DVB-DSNG standard.[1]

Synchronization and detection issues

The problem of digital synchronization for TC-QAM has been covered in the past,[8].[6] In particular, the latter reference provides a pragmatic solution to the problem of timing, amplitude and phase estimation for TC-16QAM signals. The approach followed was to use a robust Non Data-Aided (NDA) scheme proposed by Gardner[13] for timing recovery. For amplitude estimation a NDA algorithm is used for the acquisition phase, switching to a Decision-Directed (DD) scheme after lock. The same approach described in[6] has been exploited in the following.

Performance results

In this section we summarize results of computer simulations performed for the trellis coded 16-APSK constellations studied in the previous sections and compared to 16-QAM. First, AWGN linear channel results are discussed while simulations for the satellite nonlinear channel will be reported in a following sub-section. The reference system utilized in the simulations has been discussed previously.

When not stated otherwise we assume that the signal is band-limited in transmission with a square-root raised-cosine filter with roll-off factor 0.35 as for the current DVB-S standard.[1] All the following results include syn-

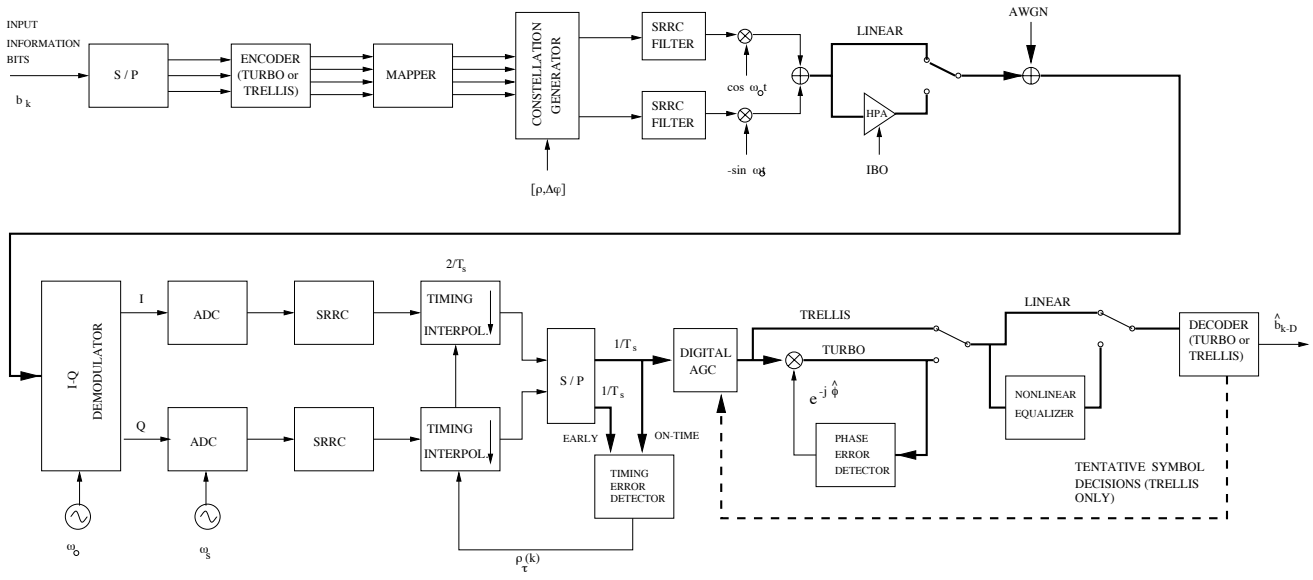

Fig. 6 End-to-end system block diagram.

chronization losses [1].

In Fig. 7, simulation results for the TCM case are reported, with the following configurations: 1) "Classical" 16-QAM with linear equalization (By "classical" we mean that the receiver does not take into account the distortion in the received constellation for demodulation and decoding.), 2) "Classical" 16-QAM with nonlinear equalization (A third order 30 taps Volterra nonlinear equalizer, with linear adaptation step $\alpha_{\text{lin}} = 10^{-3}$ and a nonlinear adaptation step $\alpha_{\text{nonlin}} = 10^{-6}$, has been simulated.),[16] 3) 4+12-PSK with linear equalization, 4) 4+12-PSK with nonlinear equalization,[16] 5) 16-QAM with linear equalization and modified TC decoder metrics,[6] 6) 16-QAM with nonlinear equalization[16] and modified TC decoder metrics,[6] 7) 4+12-PSK with pre-compensation and linear equalization, 8) 4+12-PSK with pre-compensation and nonlinear equalization.[16] In the nonlinear channel, results demonstrate the superior performance of the proposed double-ring PSK modulation. Notice, moreover, that at the BER of interest, the performance gap between the optimal pre-compensated 4+12-PSK and its non-optimal counterpart, is much less than that for optimal and non-optimal post-compensated 16-QAM, demonstrating once again the inherent robustness of the proposed schemes over nonlinear channels. At the same time while pre-compensation for 4+12-PSK can still provide about 0.5 dB performance improvement, equalization becomes totally superfluous, reducing the necessary demodulator complexity.

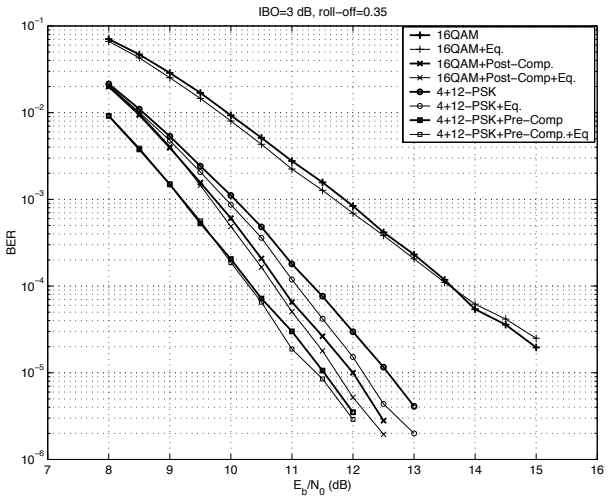

Fig. 7 Simulated BER for TCM in the nonlinear channel: r=3/4 code, IBO= 3 dB, roll-off factor=0.35.

TURBO-CODED PERFORMANCE

Overall System Description

The end-to-end system under consideration is described in Fig. 6. The modulator part is identical to the TCM one described in the coding section only the TCM coder is replaced by a rate $r = 3/4$ turbo encoder. Also the demodulator down to the baseband serial-to-parallel symbol samples converter is identical to the TCM one. Both S/P streams are used to derive the timing error information.

[1]The following demodulator synchronizer settings have been adopted throughout this section: AGC adaptation step $\gamma_\alpha = 10^{-3}$, symbol clock estimator loop noise bandwidth $B_{L\tau} = 10^{-3}$ and phase estimator loop noise bandwidth $B_{L\theta} = 5 \cdot 10^{-4}$.

The on-time single sample per symbol then enters the digital AGC described in turbo synchronization section. The digital AGC performs only amplitude gain adjustment. A dedicated phase error detector is present at the output of the digital AGC as previously discussed. If required, QAM frequency estimators are applicable to our case.[12] The turbo decoder represents the last demodulator block.

Code selection and complexity considerations

This section covers the impact that the use of turbo codes has on the overall transmission system. In particular, it describes in some detail the relevant trade-offs for the code selection, the modulation mapping choice, and the decoder operation.

The basic question to be answered for code selection is how to combine coding and modulation. The classical approach to this problem, as originally championed by Ungerboeck,[8] and used in practice in many systems, such as the one mentioned in the Introduction, is to combine coding and modulation in a single entity, namely+ Coded Modulation. The previously mentioned TCM is a good example of such a combination. At the receiver side, the demodulation and decoding are also operated jointly. Several attempts at following this road have been undertaken in the turbo coding community,[4] with varying degrees of success. In general, the obtained performance is good or even very good, as can be expected from the presence of turbo codes. However, the receiver complexity easily exceeds the practical limits, and most importantly, the simplicity and elegance of the original Ungerboeck approach gets lost in a sea of ad-hoc adaptations.

A somewhat more radical possibility is that based on the so-called Bit-Interleaved Coded Modulation (BICM), formalized by Caire et al.,[19] and commented and expanded in more detail elsewhere.[17] This new coded modulation general scheme drops one of the axioms of Ungerboeck's paradigm, that of combining encoding and modulation in a single entity. It can be shown[19] that the theoretical loss in capacity by doing so is minimal if Gray mapping is used. A key innovation of these schemes, however, is that of realizing that the receiver should operate jointly the demodulation and the decoding parts. This is, in particular, the way the TCM is operated, at least from the receiver side (let us recall that the modulation symbols are also code symbols, which means in turn that the decoding process can work on the reliabilities of the channel symbols directly, without any loss in performance). In many cases, the code symbols cannot be mapped directly on to modulation symbols, which can be expected to degrade performance. BICM solves this problem elegantly thanks to the presence of a bit-level interleaver between the two blocks, allowing the receiver to operate in a quasi-optimal way, even though a demodulation-to-bit step is introduced. A last point is that the code can be selected as a good binary code, without any further optimization or search similar to that for codes optimized for set partitioning mapping.

The road we have followed for the Coded Modulation selection is clear, namely, we have chosen a BICM scheme, in which a good binary code is mapped onto the non-binary modulation symbols through Gray mapping (see Fig. 8).

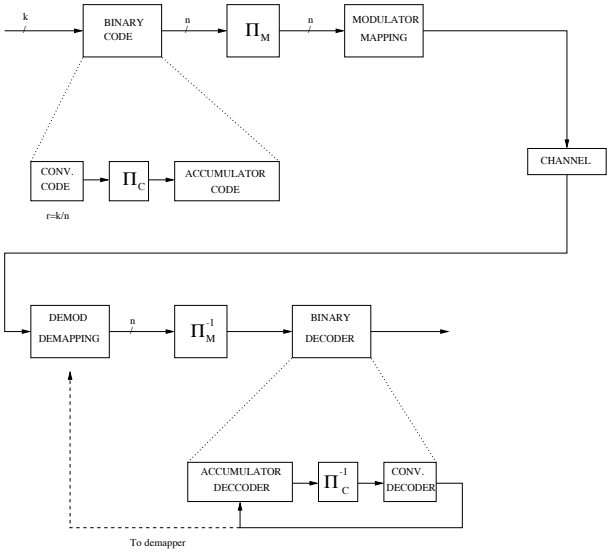

Fig. 8 BICM block diagram.

After this point, we discuss the selection of the binary code itself, based on lessons learned from the developments in turbo coding since their invention. The key point is the simple, yet fundamental, fact that a code which performs very well can be obtained from a not-too-complex good binary code. The output of this code, taken as a block, is randomly permuted, and then passed through a simple accumulator code.[18] In[17] it is shown that this procedure can improve the performance, and approach the ultimate Shannon limit by adding a certain number of these permutation-accumulator blocks at the output of the first code. As a matter of fact, the first block could be as simple as a repetition code, although some decoder implementation problems prevent us from choosing such a simple scheme. An high-level coder and decoder block diagram is shown in Fig. 8.

Decoding of this concantenated code proceeds iteratively, as is done for turbo codes. The demodulator output is transformed into metrics, the code is decoded serially,

and then, information is fed back to the input. This feedback can be to the de-mapper or at the accumulation code, as shown in Fig. 8. This procedure is repeated for a certain number of iterations, either a fixed value or until convergence is detected by means of some stopping rule.

Taking into account these remarks, two codes of rate 3/4 have been chosen for a final trade-off. The first is based on the optimum 3/4 convolutional code, with 16 states. The second is in turn built around the optimum 64-state, rate 1/2 convolutional code, punctured to rate 3/4. Although both have similar performance in the so-called "waterfall" region, where the former beats the latter by about 0.15 dB, the second is expected to have a lower floor, well below the QEF point. The whole coding and modulation subsystem consists of the serial concatenation of this code, connected with an accumulator through a pseudo-random permutation. The mapping to the modulation is the already mentioned Gray mapping, with a bit-interleaver between the stages of coding and modulation.

Concerning the decoding complexities of these codes, a simple calculation gives an increase in complexity (with respect to a soft Viterbi decoder for the standard 64-state, rate 1/2 convolutional code), of about 20-40 times for each case. The decoding algorithm for each constituent code, including the accumulator, is assumed to be 4 times as complex as the Viterbi algorithm would be for the same code. The decoding is done for a total of 10 iterations.

The remaining point to be considered is the decoding algorithm. The first step to be performed is the computation of the reliability of each received symbol. This is basically a calculation of the *a posteriori probability* $p(a_k|x_k)$ of each modulation symbol a_k, taking into account the received complex sample x_k. This is closely related to the computation of the geometric distance between the demodulator complex samples and the reference 4+12-PSK constellation. The AGC and synchronization subsystems ensure a good match between the replica and the demodulator output samples. The previous step is done without any knowledge of the code itself. The decoding procedure consists in updating the *a posteriori* probabilities $p(b_k|\mathbf{x})$ of the information bits b_k, with $\mathbf{x}$ is the whole received sequence of complex symbols, taking into account all the received sequences $p(a_k|x_k), \forall k$, and the code structure. This is achieved through the so-called iterative sum-product decoding of the concatenated code, an approach introduced for turbo codes.[11] A detailed presentation of the algorithm is outside the scope of this paper. Again in [17] it is shown numerically that for a small number of permutation/accumulator blocks, the practical performance is good enough, within 1 dB of the Shannon limit for usual operating points.

Synchronization and demodulation issues

The symbol clock estimation devised for trellis-coded modulation is also fully applicable to the current case of turbo coding. In fact the timing recovery algorithm described in[6] works without any data knowledge and performs well even at low signal to noise ratios. Concerning the AGC, in the case of a turbo coded signal we adopted the non data aided (NDA) approach described for trellis-coded modulation during the acquisition phase and reported in.[6] For a loop adaptation step $\gamma_\alpha = 5 \cdot 10^{-4}$, smaller than for TCM, the amplitude rms error provided by the NDA AGC, although higher than for TCM, is fully acceptable. In the case of turbo coding, the DD scheme exploiting trellis decoder tentative decisions (see[6]) cannot be adopted because of the unacceptable decoding delay which will cause instabilities in the carrier phase estimation loop. An approach pursued in[21] consists in the exploitation of the decisions of the first convolutional decoder and the hard decisions for the remaining coded symbols to wipe-off data modulation in the phase estimation process. However, the symbol decisions provided in this way are not expected to be significantly better than hard-decisions at the decoder input considering the very low operating SNR typical of a turbo decoder, and the weakness of the constituent codes.

In our specific case, the 16-ary APSK modulation symbol decisions required to wipe-out data modulation effects may be replaced by a simpler algorithm based on a four quadrant phase averaging,[5] as illustrated in Fig. 9. The idea consists in processing the complex symbol matched filter samples according to the quadrant to which they belong. This corresponds to a simplified (amplitude independent) decision directed scheme, in which the original M-ary constellation collapses to the four points: $(\pm 1 \pm j)/\sqrt{2}$. In this way the average signal carrier phase can be simply estimated and the scheme is also robust to amplitude errors, thus easing the initial acquisition transient. The proposed scheme averages out the signal points falling in each of the four complex plane quadrants exploiting the constellation symmetry to derive the reference phase with an ambiguity equal to $2\pi/12$. The proposed NDA phase error estimator has the advantage of being insensitive to possible amplitude errors that are typical of the acquisition phase. The proposed scheme corresponds to a DD phase estimator for

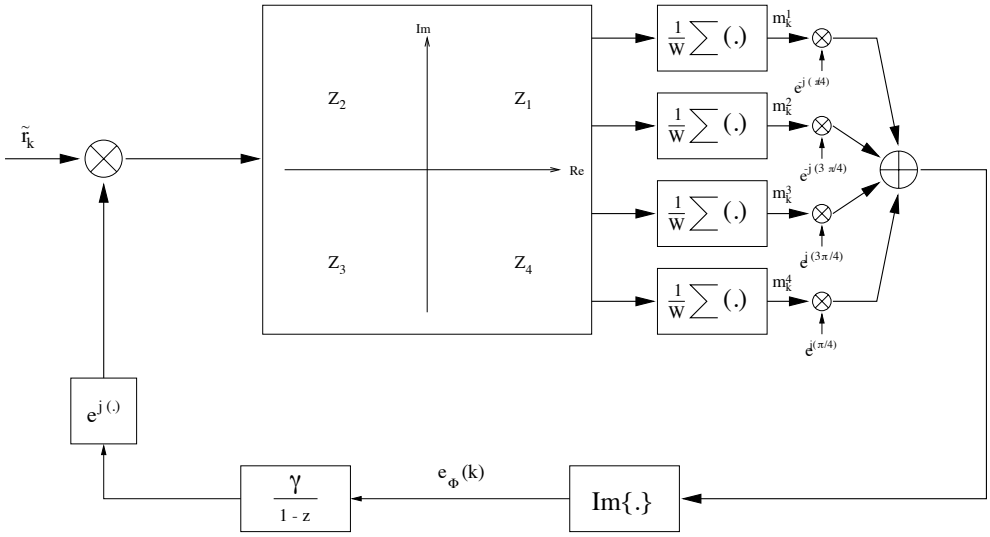

Fig. 9 Four Quadrant Phase Estimator.

QPSK.[22]

Because of the simplified four-quadrant decision approach its implementation is easier than a truly DD scheme for 4+12-PSK and decisions are insensitive to possible signal amplitude estimation errors. The main scheme drawback is related the so called pattern noise or irreducible phase jitter floor due to the phase error signal averaging for the four constellation points belonging to the same quadrant. This floor is not present with QPSK modulation, where the exact Decision-Directed, and this simplified version are identical. Simulation results indicate that the performance is degraded with respect to the TC-DD scheme but remains acceptable even at low SNR. It was found that a phase jitter standard deviation of about 1 degree can be achieved at $E_b/N_0 = 5$ dB for loop bandwidth[12] $B_{L\Phi}T_s(\infty) = 5 \cdot 10^{-5}$.

Performance results

We will now report some performance results obtained exploiting the selected 4+12-PSK modulation format and turbo coding scheme. The code used for the simulations was the one based on the rate 3/4 16-state convolutional code, although the results are similar for the two cases studied in the turbo code section. Although we have not verified the results down to Quasi Error Free (QEF) operation, we have confidence in the fact that QEF is achieved at values close to the ones mentioned in the paper. This is due to the fact that the error floor asymptote is very low, particularly for the 64-state code, and that the slope is very steep, down to the points we have simulated. It is therefore safe to extrapolate the performance for QEF from the values simulated in practice.

The following demodulator synchronizer settings have been adopted throughout this section: an AGC loop adaptation step of $\gamma_\alpha = 10^{-3}$, the one-sided loop noise bandwidth of the symbol clock estimator, $B_{L\tau} = 10^{-3}$ and the one-sided loop noise bandwidth of the carrier phase estimator $B_{L\Phi} = 2 \cdot 10^{-5}$. As for trellis coded modulation we used a third-order timing interpolator. For the simulations the rate 3/4 turbo codec described previously has been adopted with an input frame block size of 16384 bits. When not specified, the number of decoder iterations has been set to 10. Results for the AWGN channel are provided with and without channel estimation in Fig. 10. We observe that for QEF performance over the linear channel the turbo coding gain is about 1.65 dB with respect to the 16 state trellis code concatened with a trellis decoder and only 1.15 dB with respect to the 32-state trellis code.[5] Additionally, we should add the 0.35 dB loss due to the Reed Solomon concatenation for the TCM case, absent for the turbo as mentioned before. This amounts to a total gain of about 2.0 dB, for an increase in complexity of a factor 20-40. For the sake of comparison, the same figure shows the performance of 16-QAM with the same turbo code as well. The impact of the digital synchronizer described in a previous section has been verified by simulation and it has been found to be less than 0.1 dB in the waterfall BER region of the turbo decoder. This was considered a very satisfactory result considering the simplicity of the NDA scheme adopted. Considering the limited gain provided by the nonlinear equalizer for trellis-coded modulation, we abandoned the scheme for the turbo case. The IBO was optimized, as detailed in.[5] The E_b/N_0 was selected in the

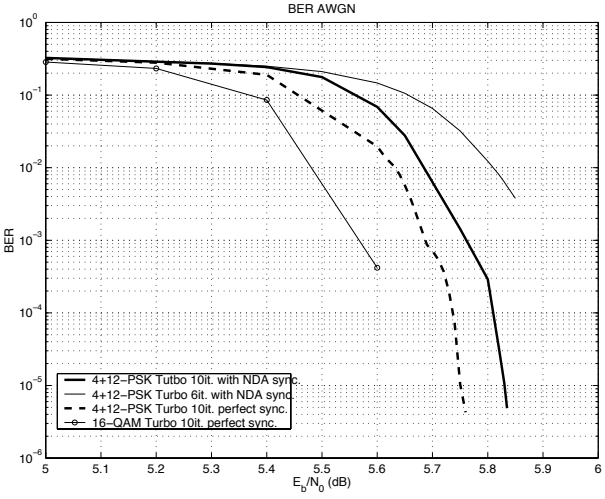

Fig. 10 Simulated BER for 16-QAM and 4+12-PSK in the AWGN channel: rate 3/4 turbo code.

middle of the waterfall BER region to allow for capturing the BER variation. Clearly, optimization cannot be performed at $P_b = 10^{-10}$ but the current working point hardly differs from the QEF in terms of SNR. The decoder input SNR was computed based on the effective demodulator input E_b/N_0 previously defined (see eqn. (6)) $[E_b/N_0]_{\text{eff}}$. For the optimum IBO of 2 dB the BER over the nonlinear channel was then simulated and the results are collected in Fig. 11. We can see that due to the optimal combination of

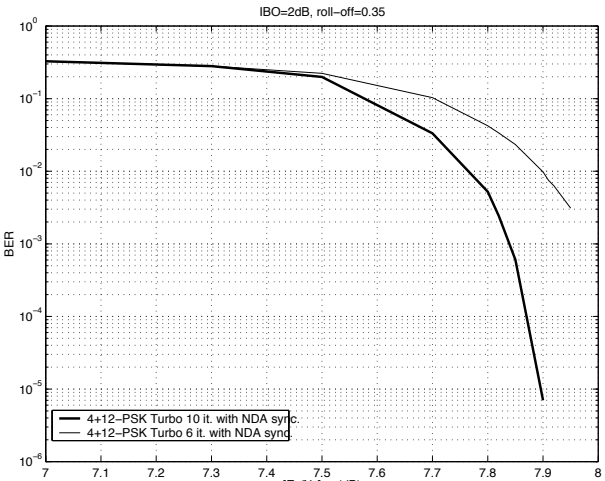

Fig. 11 Simulated BER for turbo coded 4+12-PSK in the nonlinear channel: r=3/4 code, IBO= 2 dB, roll-off factor=0.35.

a powerful turbo code and a robust 4+12-PSK modulation we can operate over a typical satellite nonlinear channel for QEF performance at only 1.3 dB OBO with a degradation with respect to the linear channel of only 0.9 dB and an operating $[E_b/N_0]_{\text{inp}} = 6.9$ dB or $[E_b/N_0]_{\text{sat}} = 8.2$ dB. This shall be compared to the $[E_b/N_0]_{\text{sat}} = 13.1$ dB required by conventional TC-16-QAM over nonlinear satellite channels.

This represent an improvement of about 5 dB in power and 8 % in spectral efficiency compared to the conventional TC-16-QAM representing today's baseline for high-speed satellite links.

SUMMARY AND CONCLUSIONS

In this paper a new class of modulation, called double ring 16-ary modulations, has been devised. It has been shown that over AWGN channels this kind of constellation is very close to, and may even slightly outperform the capacity of classical coded 16-QAM. For nonlinear satellite channels the so-called 4+12-PSK double ring constellation has been selected, as the robustness to nonlinear effects compensated for a minor loss in performance with respect to 16-QAM over AWGN linear channels. Amplitude, carrier phase and symbol clock recovery schemes have been devised for both decision-directed and non data-aided operations. Simple yet efficient pre-distortion schemes have been studied. Results indicated that compared to TC-16QAM currently adopted for high-speed satellite links with linear equalization the proposed 4+12-PSK with pre-distortion and Turbo coding can achieve a gain superior to 5 dB in power efficiency and 8 % improvement in spectrum efficiency.

References

[1]M. Cominetti and A. Morello, *"Digital Video Broadcasting over satellite (DVB-S): a system for broadcasting and contribution applications"*, Int. Jour. on Satellite Commun., 2000, No. 18, pp. 393-410.

[2]J. Farserotu, R. Prasad, *"A Survey of Future Broadband Multimedia Systems, Issues and Trends,"* IEEE Comm. Magazine, June 2000, pp. 128-133.

[3]S. Benedetto, E. Biglieri and V. Castellani, "Digital Transmission Theory", Prentice Hall, 1987.

[4]D.J.Costello, Jr. J. Hagenauer, H. Imai, S. B. Wicker,*"Applications of Error Control Coding"*, IEEE Trans. on Inform. Theory, vol 44 pp. 2531-2560, October 1998.

[5]R. De Gaudenzi, A. Guillén i Fàbregas, A. Martinez Vicente, B. Ponticelli,*"A New Coded Digital Modulation Scheme for Nonlinear Satellite Channels with High Power and Spectral Efficiency,"* ESA Technical Report, ESA STR-242, July 2001.

[6]R. De Gaudenzi and M. Luise, *"Design and Analysis of an All-Digital Demodulator for Trellis Coded 16-QAM Transmission over a Nonlinear Satellite Channel,"* IEEE Trans. on Comm., Vol. 43, No. 2/3/4, February/March/April 1995, part I.

[7]G. F. Forney Jr., R. G. Gallager, G. R. Lang, F. M. Longstaff, S. U. Quereshi, *"Efficient Modulations for Band-Limited Channels,"* IEEE Journ. on Sel. Areas in Comm. Vol. SAC-2, No. 5, Sept. 1984, pp. 632-647.

[8]G. Ungerboeck, *"Channel Coding with Multilevel Phase Signals,"* IEEE Trans. on Information Theory, Vol. IT-28, January 1982.

[9]M. Rouanne and D. J. Costello, *"An Algorithm for Computing the Distance Spectrum of Trellis Codes,"* IEEE Jour. on Sel. Areas in Comm., Vol. 7, N. 6, August 1989.

[10]S. Pietrobon, R. H. Deng, A. Lafanechere, G. Ungerboeck, D. J. Costello Jr., *"TC Multidimensional Phase Modulation,"* IEEE Trans. on Information Theory, Vol. 36, N.1, January 1990.

[11]G.F. Forney Jr., *"Codes on Graphs: Normal Realizations,"* IEEE Trans. on Information Theory, Vol. IT-35, February 2000.

[12]U. Mengali and A. N. D'Andrea, "Synchronization Techniques for Digital Receivers," Plenum Press New York, 1997.

[13]F. M. Gardner, *"A BPSK/QPSK Timing-Error Detector for Sampled Receivers,"* IEEE Trans. on Commun., Vol. COM-34, N. 5, May 1986.

[14]F. M. Gardner, *"Interpolation in digital modems - Part I: Fundamentals"*, IEEE Trans. on Comm., Vol. 41, No. 3, March 1993.

[15]L. Erup, F. M. Gardner, R.A. Harris *"Interpolation in digital modems - Part II: Implementation and performance"*, IEEE Trans. on Comm., Vol. 41, No. 6, June 1993.

[16]S. Benedetto, E. Biglieri, *"Nonlinear Equalization of Digital Satellite Channels,"*, IEEE Jour. on Sel. Areas in Comm., Vol. SAC-1, No. 1, Jan. 1983, pp. 57-62.

[17]A. Martinez, B. Ponticelli, *"Efficient coding schemes for high order modulations in satellite broadcasting applications,"* Proc. of the 7th International ESA Workshop on Digital Signal Processing Techniques for Space Communications, Lisbon, October 2001.

[18]H. Jin, R.J. McEliece *"RA Codes Achieve AWGN Channel Capacity,"* Proc. 13th Int. Symp. 3th International Symposium on Applied Algebra, Algebraic Algorithms, and Error-Correcting Codes, AAECC-13, Honolulu, Hawaii, November 1999.

[19]G. Caire, G. Taricco, E. Biglieri, *"Bit Interleaved Coded Modulation,"* IEEE Trans. on Information Theory, vol. 44, pp. 927-947, May 1998.

[20]*"TOPSIM-IV Users Manual,"* Politecnico di Torino, Italy, January 1992.

[21]C. Langlais, M. Hèlard, *"Phase Carrier Recovery for Turbo codes over a satellite link with the help of tentative decisions,"* In the proc. of the 2^{nd} International Symposium on Turbo codes and related topics, Brest, France, 2000, pp. 439-442.

[22]R. De Gaudenzi, T. Garde, V. Vanghi, *"Performance Analysis of Decision-Directed Maximum-Likelihood Phase Estimator for M-PSK Signals,"* IEEE Trans. on Comm., Vol. 43, No 12, December 1995.