Performance analysis of turbo-coded APSK modulations over nonlinear satellite channels

Citation for published version (APA):

DOI:
10.1109/TWC.2006.1687763

Document status and date:
Published: 01/01/2006

Document Version:
Publisher’s PDF, also known as Version of Record (includes final page, issue and volume numbers)

Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher’s website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

Link to publication

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the “Taverne” license above, please follow below link for the End User Agreement:
www.tue.nl/taverne

Take down policy
If you believe that this document breaches copyright please contact us at:
openaccess@tue.nl
providing details and we will investigate your claim.

Download date: 17. Sep. 2023
Abstract—This paper investigates the performance of \( M \)-ary Amplitude-Phase Shift Keying (APSK) digital modulation over typical nonlinear satellite channels. The effect of the satellite nonlinearity is studied, and distortion pre- and post-compensation techniques for coded APSK are presented. Moreover, clock timing, signal amplitude and carrier phase recovery schemes are discussed. For the latter, a new class of non turbo decoder-aided closed-loop phase synchronizers featuring good performance and low complexity is studied. Finally, an end-to-end coded APSK system simulator inclusive of the satellite channel model and synchronization sub-systems is discussed and its performance compared to standard trellis-coded QAM concatenated with Reed-Solomon codes, showing a remarkable gain in both power and spectral efficiency. Coded APSK, recently selected for the new standard –DVB-S2– for digital video broadcasting and interactive broadband satellite services [1], is shown to represent a power- and spectral-efficient solution for satellite nonlinear channels.

Index Terms—Turbo codes, amplitude-phase shift keying (APSK) modulation, bit-interleaved coded modulation (BICM), coded modulation, nonlinear channels, phase synchronization, time synchronization, satellite communications.

I. INTRODUCTION

The strength of satellite communication systems lies in their ability to efficiently broadcast digital multimedia information over very large areas [2]. 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 multibeam satellite antennas with adaptive coding and modulation (ACM) schemes will allow an important efficiency increase for satellite systems operating at Ku- or Ka-band [3]. As a result, to maximize the efficient use of the scarce on-board available power, the satellite high power amplifier (HPA) must be operated close to saturation, creating a highly nonlinear environment, in which efficient transmission of more than 3 bps/Hz is a challenging task.

Past literature has extensively covered techniques to improve highly spectral efficient modulations performance [4]–[7]. However, this work typically refers to uncoded square QAM constellations, results of lesser importance for nonlinear satellite channels. Circular APSK modulations were already proposed thirty years ago in [8], where uncoded bit error rate bounds for several non band-limited APSK sets were provided; 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 revert this conclusion.

The possibility of modulator pre-compensation was briefly mentioned in [8]. General techniques to mitigate the nonlinear distortion effects are well covered in the literature [4], [5], [7], [9]–[15]. Pre-compensation techniques counteract the amplifier distortion through constellation pre-distortion at the transmitter, while post-compensation techniques mitigate nonlinear distortion effects at the demodulator side through nonlinear equalization and/or ad-hoc decoder metric computation. In this paper we review the existing works and we discuss their applicability to APSK.

We also consider clock timing, amplitude and phase recovery schemes, with the final aim of designing an all-digital low-complexity high performance demodulator for coded APSK modulations. In [14] an all-digital trellis coded 16-QAM demodulator was studied. In that paper nonlinearity compensation was limited to the demodulator trellis decoder, whose metric calculation was modified to take into account the major satellite memoryless nonlinear distortion effects. Joint iterative decoding and parameter estimation can be performed by adding the parameter variables in the message-passing iterative decoding algorithm [16], [17]. However, each parameter to be estimated in the iterative loop slows the convergence of the iterative decoder and changes its properties, thus significantly increasing the overall complexity of the demodulator. In this work, we consider low-complexity closed-loop modules for timing, amplitude and phase synchronization, with sub-optimum yet good performance.

This paper is organized as follows. Sect. II describes the system model and the main effects of the nonlinear amplifier. Sect. III briefly describes APSK signal sets and presents an optimization procedure based on mutual information. In Sect. IV we review the pre- and post-compensation techniques against nonlinear distortion, and discuss their applicability to coded APSK systems. Sect. V describes the proposed low-complexity demodulator, in particular the clock timing,
amplitude and carrier phase estimation schemes. Sect. VI gives
a brief description of the system we use as benchmark for
comparison, and shows some numerical results that illustrate
the benefits of the proposed system over today standard. Main
conclusions and findings are finally summarized in Sect. VII.

II. SYSTEM MODEL

We consider a communication system composed of a digital
modulator, a square-root raised cosine (SRRC) band-limiting
filter, and a nonlinear high power amplifier (HPA) with typical
Travelling Wave Tube Amplifier (TWTA) characteristic for
a satellite in Ku/Ka-band. This model is representative of a
satellite bent-pipe transponder with uplink noise negligible
driven at saturation by a continuous wave (CW) carrier and
shows some numerical results that illustrate the impact of the
satellite input and output analog filters is assumed negligible. The baseband equivalent of the
transmitted signal at time $t$, $s_T(t)$ is given by:

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

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 $|\mathcal{X}| = M$, described in detail shortly hereafter, $p_T(t)$
is the SRRC transmission filter impulse response, and $T_s$ is
the symbol duration (in seconds), corresponding to a channel
use. The coded modulation spectral efficiency $R$ is the number
of information bits divided by the modulator baud rate, i. e.,
$R = r \log_2 M$, $r$ being the coding rate and $M$ the modulation
cardinality.

The signal $s_T(t)$ passes through an 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:

$$s_A(t) = F(|s_T(t)|) e^{j\Phi(|s_T(t)|)}$$

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 $s_T(t) = |s_T(t)| e^{j\phi(s_T(t))}$. For the numerical examples in the following we will consider the AM/AM and AM/PM characteristics shown in Fig. 1.

We introduce the parameter $E_b/N_0|_{\text{sat}}$ 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 at the demodulator input [10]. Note that the sub-fix “sat” refers to HPA saturation. The signal-to-noise ratio at the demodulator input $E_b/N_0|_{\text{inp}}$ is reduced by the output back-off (OBO, in dB) with respect to the value in a system operating with a single constant-envelope signal at amplifier saturation:

$$\frac{E_b}{N_0|_{\text{sat}}} \text{ (IBO)} = \frac{E_b}{N_0|_{\text{inp}}} \text{ (IBO)} + \text{OBO (IBO)} \text{ (in dB)}$$

Additionally, due to constellation warping and satellite channel induced ISI, the demodulator performance is degraded by an amount $D$ with respect to an ideal linear AWGN channel. This quantity depends on the HPA distortion and hence on the IBO/OBO characteristic (see Fig. 1). With this degradation, the effective demodulator input $E_b/N_0|_{\text{eff}}$ is given by:

$$\frac{E_b}{N_0|_{\text{sat}}} \text{ (IBO)} = \frac{E_b}{N_0|_{\text{eff}}} \text{ + OBO (IBO) + D (IBO)} \text{ (in dB)}$$

With Eq. (4) we can find an optimum HPA operating point that minimizes $E_b/N_0|_{\text{sat}}$. This point represents the best trade-off between the increasing power loss (OBO) related to the higher IBO and the reduction of the distortion ($D$) due to the improved linearity experienced by a larger IBO.

III. CONSTELLATION OPTIMIZATION IN AWGN

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:

$$\mathcal{X} = \left\{ \begin{array}{ll}
 r_1 e^{j\frac{\pi}{n_1} + \theta_1} & i = 0, \ldots, n_1 - 1, \\
 r_2 e^{j\frac{\pi}{n_2} + \theta_2} & i = 0, \ldots, n_2 - 1, \\
 \vdots & \\
 r_{n_R} e^{j\frac{\pi}{n_{n_R}} + \theta_{n_R}} & i = 0, \ldots, n_{n_R} - 1,
\end{array} \right\}$$

where we have defined $n_\ell$, $r_\ell$ and $\theta_\ell$ as the number of points, the radius and the relative phase shift corresponding to the $\ell$-th ring respectively. We will name such modulations as $n_1 + \ldots + n_{n_R}$-APSK. Fig. 2 depicts the 4+12- and 4+12+16-APSK modulations with quasi-Gray mapping. In particular, for next generation broadband systems [1], [3], 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 $r_\ell$ are normalized such that $\sum_{\ell=1}^{n_R} r_\ell^2 = 1$. Notice also that the radii $r_\ell$ are ordered, so that $r_1 < \ldots < r_{n_R}$.

For convenience we define the phase shifts and the ring radii in relative terms rather than in absolute terms, as in (5). We let $\phi_\ell = \theta_\ell - \theta_1$ for $\ell = 1, \ldots, n_R$ be the phase shift of
the \( \ell \)-th ring with respect to the inner ring. We also define \( \rho_\ell = r_\ell/r_1 \) for \( \ell = 1, \ldots, n_R \) as the relative radii of the \( \ell \)-th ring with respect to \( r_1 \). In particular, \( \phi_1 = 0 \) and \( \rho_1 = 1 \).

We are interested in finding an APSK constellation, defined by the parameters \( \mathbf{\rho} = (\rho_1, \ldots, \rho_{n_R}) \) and \( \mathbf{\phi} = (\phi_1, \ldots, \phi_{n_R}) \), such that a given cost function \( f(\mathbf{X}) \) reaches a maximum. Rather than the classical minimum Euclidean distance, mutual information allows us to optimize the performance at the coded system \( E_b/N_0 \) operating point. We assume a linear AWGN channel; robustness against nonlinear distortion will be achieved in a second step through exploitation of constellation pre-compensation.

The 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. [18]),

\[
f(X) = I(X; Y) = \log_2 M - \frac{1}{M} \sum_{x \in \mathcal{X}} \mathbb{E}_n \left[ \log_2 \left( \sum_{x' \in \mathcal{X}} e^{-\frac{1}{n_0} \sqrt{\mathcal{E}_n(x-x') + n |n|^2}} \right) \right].
\]  

(6)

It was found in [9] that the maximum mutual information does not depend on \( \phi \). Therefore, the optimization can be done by simply finding the \( \rho_2 \) that maximizes the mutual information. This result was found to hold true also for the other constellations and hence, in the following, mutual information optimization results do not account for \( \phi \). Fig. 3 compares the mutual information of optimized 16-APSK modulations versus 16-PSK and 16-QAM. Table I summarizes the optimized 16- and 32-APSK parameters for various coding rates. More details on constellation optimization and code design can be found in [19]. These values are used in the DVB-S2 standard [1].

IV. SATELLITE CHANNEL DISTORTION COMPENSATION

Generally speaking, the amplifier non-linearity has two major effects. First, the very constellation is distorted as the constellation points are mapped by the HPA nonlinear filter. This leads to an overall nonlinear channel with memory. As a consequence, the demodulator SRRC is not matched anymore to the incoming signal. This issue is to be tackled mainly with a pre-equalization at the modulator [5], or equalization at the demodulator or a combination of the two techniques (see Sect. IV-A.2-IV-B).
compensation parameters can be calculated off-line, in the following steps: 1) Generation of $S$ blocks of $W$ symbols over which the symbol matched filter (SMF) centroids are computed (transmission in the absence of white noise); 2) Computation of the error signal at the end of each block; 3) Pre-distorted constellation point update. The latter task can be readily achieved through an iterative least mean square (LMS) type of algorithm illustrated by the following equations:

$$|x_{\text{pre}}^{(n)}(s+1)| = |x_{\text{pre}}^{(n)}(s)| - \gamma_t e_c^{(n)}(s)$$  \hspace{1cm} (7)$$

$$\arg(x_{\text{pre}}^{(n)}(s+1)) = \arg(x_{\text{pre}}^{(n)}(s)) - \gamma_\phi \psi(s)$$  \hspace{1cm} (8)$$

$$r^{(n)}_c(s) e^{j\theta^{(n)}_c(s)} = \frac{1}{W} \sum_{k \in \{n\}} y(k)$$  \hspace{1cm} (9)$$

$$\psi(s) = \begin{cases} \arg(e_c^{(n)}(s)) - 2\pi & \text{if } \arg(e_c^{(n)}(s)) > \pi \\ \arg(e_c^{(n)}(s)) + 2\pi & \text{if } \arg(e_c^{(n)}(s)) < -\pi \\ \arg(e_c^{(n)}(s)) & \text{if } |\arg(e_c^{(n)}(s))| \leq \pi \end{cases}$$  \hspace{1cm} (10)$$

where the index $n$ refers to the constellation point, $t^{(n)}$ indicates the conditioning to the constellation point $n$, $s$ refers to the iteration step of the algorithm, $y(k)$ represents the $k$-th SMF output complex sample, $x^{(n)}$ represents the APSK complex constellation reference point, $r^{(n)}_c(s)$ and $\theta^{(n)}_c(s)$ are the modulus and the phase of the SMF output complex $n$-th centroid computed at step $s$, and $x_{\text{pre}}^{(n)}(s)$ the pre-distorted $n$-th constellation point computed at step $s$, $\gamma_t$ and $\gamma_\phi$ the adaptation steps for the pre-distorted constellation point modulus and the phase respectively.

For example, the optimal 4+12-APSK pre-distortion parameters are $\rho_2' = 3.5$ and $\Delta \phi = 25$ deg for an IBO=3 dB and $\rho_2' = 3.7$ and $\Delta \phi = 27$ deg for a smaller IBO=2 dB (see Fig. 5-a). As expected, the pre-distorted constellation is expanded, e. g., $\rho_2 > \rho_2'$. The new constellation points $x'$ follow (5), 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 effect of the HPA on the phase HPA between inner and outer rings through a simple change in 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 $r'$, $\phi'$, with no hardware complexity impact or out-of-band emission increase at the linear modulator output. The compensation effort is shifted into the modulator side, allowing the use of an optimal demodulator/decoder for AWGN channels even when the amplifier is close to saturation. The signal at the modulator output is then

$$s_{T}^{\text{pre}} = \sqrt{P} \sum_{k=0}^{L-1} x'(k) p_T(t - kT_s)$$  \hspace{1cm} (12)$$

where now $x'(k) \in X'$ being the pre-distorted symbols with $r'_1$ and $\phi'_1$ for $t = 1, \ldots, n_R$.

To show the effect of this compensation, the scatter diagram at the output of the SRRC (with roll-off factor 0.35)
receiver matched filter is shown in Fig. 5-b for 16-APSK with IBO=2 dB. For clarity, the scatter diagram at the SMF has been obtained in the absence of AWGN. Despite the strong channel nonlinearity the center of mass, corresponding to the scattered diagram, closely follows the optimum 4+12-APSK constellation, for which the optimum parameters are $\rho_2 = 2.7$ and $\Delta \phi = 27$ degrees.

Measurements showed that the HPA characteristic sensitivity to temperature or aging results in a limited change of gain but not in a modification of the AM/AM, AM/PM characteristics shape. The limited gain variations are compensated by the satellite transponder automatic level control (ALC) device, thus off-line pre-compensation has a long term value. If required, the compensated parameters can be adapted to track larger slow variations in HPA characteristic due to aging.

2) Dynamic Distortion Compensation: The dynamic pre-compensation approach by Sari [5] considers not only the current constellation symbol, as done in the case of static pre-compensation, but also for $Q$ symbols before and after. In general for M-QAM modulation the possible pre-distortions are now $M^{2Q+1}$. Similar to the static pre-compensation described before, the pre-distorter complex values can be obtained offline minimizing the minimum squared error (MSE) between the ideal constellation and the noiseless points measured after the demodulator symbol matched filter. This can be achieved through an extension of the methodology described for the static approach. In this case, the number of updating equations for M-QAM is $M^{2Q+1}$ which can be reduced to $M^{2Q+1}/4$ exploiting the M-QAM quadrant symmetry [5]. For APSK the memory requirement can be further reduced to $3M^Q/16$.

The main drawback of this technique is the amount of time required to compute the pre-distortion coefficients and the memory required which grows rapidly with the constellation and memory parameters ($M$, $Q$). It should also remarked that the dynamic pre-compensation is less effective than the static one when more than one carrier is passing through the same HPA. This is because the uncorrelated intermodulation noise among carriers will dominate over the ISI thus making the dynamic pre-compensation ineffective.

B. Dynamic Post-compensation at the Receiver

An alternative approach to dynamic pre-distortion techniques is represented by demodulator equalization. Two main approaches are revisited in the following, namely linear and nonlinear equalization. The typical linear equalization corresponds to the case of a decision-directed complex AGC (also called vector tracker) see Appendix II of [14]. It was observed in [10] that linear equalization does not provide any significant improvement over nonlinear channels.

The post-compensation in the (trellis) decoder metric calculation was proposed in [14]. In that work, the metric computer for the TCM decoder used the distorted constellation $X'$ as reference constellation, and the metric corresponding to the distorted constellation point $x' \in X'$ is given by,

$$M(y|x') \propto \exp \left( \frac{-1}{N_0} |y - \sqrt{E_{\text{in}}} x'|^2 \right)$$

where $\propto$ indicates proportionality and the $x' \in X'$ are the centroids of the distorted constellation at the matched filter output.

A somewhat more complex approach is adaptive nonlinear equalization, based on a Volterra series expansion [10], where it is shown to provide considerable advantages for PSK modulations in terms of ISI reduction. Furthermore, it also allows to reduce the optimum operating IBO, thus improving the HPA utilization. However, it is expected that this advantage will be more pronounced for 16-QAM than for double-ring
APSK with pre-compensation, as this case is already very close to the optimal linear channel performance. A second drawback is that the improvement will be somewhat masked by the low SNR typical of powerful encoded transmission. This conjecture about the limited improvement provided by a non-linear equalizer for heavily coded modulation will be substantiated in the following (see Sect. VI).

V. Demodulation and Synchronization Issues

In this section we will describe demodulation and synchronization issues concerning the design of a complete digital demodulator for the coded APSK system. In particular we will investigate simple timing, amplitude and phase estimation algorithms able to operate at the low SNRs corresponding to a near-Shannon coded modulation.

A. Timing Recovery

The Non Data-Aided (NDA) symbol clock estimation scheme proposed by Gardner [20] can be applied, since it works without any data knowledge and exhibits very satisfactory performance even at low signal-to-noise ratios. Extraction of the k-th normalized sampling epoch $\epsilon_k$ is performed recursively via the following “closed-loop” equation,

$$\varepsilon_r(k+1) = \varepsilon_r(k) - \gamma_r e_r(k),$$  

(14)

where $e_r(k)$ is the Timing Error Detector (TED) function

$$e_r(k) = \text{Re}\left\{y(k)^*\left[y(k+\frac{1}{2}) - y(k-\frac{1}{2})\right]\right\},$$  

(15)

and $y(k)$ are the symbol matched filter output samples. As is apparent, the error detector signal is obtained using the on-time interpolated sample $y(k) \triangleq y[(k + \varepsilon_k)T_s]$ and the late/early samples $y(k+\frac{1}{2})$ and $z(k-\frac{1}{2})$ shifted by half the symbol time. The updating step-size $\gamma_r$ is related to the equivalent one-sided loop noise bandwidth $B_{Lr}$ as follows,

$$\gamma_r = \frac{4B_{Lr}T_s}{A_r(1+2B_{Lr}T_s)},$$  

(16)

where $A_r$ is the slope at the equilibrium point of the average TED characteristic (S-curve). As observed in [20], the timing discriminator performance is invariant to carrier phase, so that timing information can be extracted without prior acquisition of the carrier recovery loop. As for the digital timing interpolator, we employed the cubic interpolator described in [21] and [22].

B. Amplitude Recovery

The NDA Automatic Gain Control (AGC) algorithm for the acquisition phase described in [14] is used, for which the the $(k+1)$-th value of the variable gain $C_{k+1}$ is derived recursively as:

$$C_{k+1} = C_k - \gamma_a (C_k |y(k)| - A_1),$$  

(17)

where $A_1 \triangleq E\{|y(k)|\}$. For a loop adaptation step $\gamma_a$ of $5 \cdot 10^{-4}$, the amplitude rms error provided by the NDA AGC, is fully acceptable (see [14] for details).

C. Phase Recovery

Standard decision directed (DD) schemes exploiting trellis decoder tentative decisions, as in [14], cannot be adopted because of the long intrinsic decoding delay which will cause instabilities in the carrier phase estimation loop. The approach pursued in [23] exploits the decisions of the inner 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 inner code (accumulator, 2-state rate-1 convolutional code). In order to keep complexity within reasonable limits, and due to longer convergence delays of the joint phase detector and decoder, we do not consider iterative phase estimators as [16], [17].

In [24] we showed that a four-quadrant Decision-Directed (DD) (4Q-DD or 4Q-slicing) approach, consisting in partitioning the I-Q plane in four quadrants as for QPSK modulation and taking hard decisions at symbols level, is a workable approach for the phase recovery of 16-APSK and 16-QAM modulations. However, the estimator suffered from a phase error jitter floor due to pattern noise, an artifact absent in QPSK operation. The irreducible jitter variance for 16-APSK is computed as:

$$\sigma_{\phi,4Q-DD}^2 = \frac{1}{4W} \sum_{x \in X'} \left(\frac{\text{arg}(x') - \pi}{4}\right)^2 = \frac{1}{72W}$$  

(18)

where $X'$ represents the set of constellation points in the first quadrant of the complex plane, $x'$ is one of the elements of $X'$ and $W$ is the phase estimator averaging time in symbols.

The quadrant decision approach is somewhat fragile against thermal noise due to the vicinity of some of the 16-APSK constellation points to the decision region boundaries. The ratio between the simulated Decision Directed (DD) 16-APSK phase standard deviation and the Modulation Constrained Cramer-Rao (MCR) bound [25] is about 3.55 at $E_s/N_0 = 10$ dB, while for 16-QAM the same ratio is reduced to 1.92. This problem can be overcome by narrowing the loop bandwidth when possible. However, in the presence of important phase noise contributions originating from the consumer receiver radio frequency front-end, the narrow loop noise bandwidth reduction will generate cycle slips. Furthermore, by doing so, the maximum residual frequency error that can be tolerated by the phase recovery loop is also reduced. Thus excessive narrowing the loop noise bandwidth to combat the phase jitter should be avoided.

In the following, we propose a general class of $P$-th power closed-loop NDA phase synchronizers for APSK. The corresponding block diagram is shown in Fig. 6. Compared to the 4Q-DD scheme an $P$-th order nonlinearity followed by a $\theta_P$ phase rotation is introduced before the quadrant slicing device. The proposed phase error detector algorithm has the following form:

$$w(k) = y^P(k)e^{j\theta_P},$$  

$$e_\phi(k) = \text{Im}\left\{w(k) \cdot \left[\text{sign}(\text{Re}(w(k))) - j \text{sign}(\text{Im}(w(k)))\right]\right\}$$  

(19)
The error signal \( e_\phi(k) \) is then used to update the first order phase loop phase according to 
\[
\phi(k) = \phi(k - 1) - \gamma_\phi e_\phi(k),
\]
where the loop gain is related to the bandwidth \( B_{L,\phi} \) by the classical relation,
\[
\gamma_\phi = \frac{4B_{L,\phi}T_s}{A_\phi(1 + 2B_{L,\phi}T_s)}. \tag{20}
\]

The DD phase discriminators S-curve \( A_\phi \) depends on the actual \( E_s/N_0 \) [26]; the effective loop noise bandwidth depends in turn on the operating \( E_s/N_0 \). Following [26], it is common practice to compute the discriminator slope \( A_\phi \mid_{E_s/N_0=+\infty} \) and the corresponding loop bandwidth for \( E_s/N_0 = +\infty \), and refer in the simulations to \( B_{L,\phi}T_s(\infty) = B_{L,\phi}T_s \mid_{E_s/N_0=+\infty} \).

The 4Q-DD scheme described in [24] is obtained by setting \( P = 1 \) and \( \theta_P = 0 \) in Fig. 6. Because of the simplified 4Q-DD approach the phase estimator implementation is simpler than a truly DD scheme for 16-APSK and decisions are insensitive to possible signal amplitude estimation errors. The operations performed during this estimator suggest an alternative scheme that keeps a similar complexity, while improving its performance. Raising a 4+12-APSK constellation to the 3rd power (in the complex domain), the original 4+12-APSK constellation transforms into a QPSK one. It is easy to see that the twelve external ring points collapse into a QPSK constellation, with amplitude \( 4r_2^3 \), while the internal points remain on a QPSK constellation with amplitude \( r_1^3 \). For typical 4+12-APSK parameters the distance between the two rings is largely increased. This process is illustrated in Fig. 7, where the input 4+12-APSK signal constellation is depicted, together with the constellation after the cubic nonlinearity.

A simple 4Q-DD (quadrant slicer type) closed-loop phase estimator [26] operating on the outer QPSK ring can be efficiently applied to the cubic nonlinearity output. This scheme dubbed C-4Q-DD is obtained by setting \( P = 3 \), \( \theta_P = 0 \) in Fig. 6. The main scheme drawback is related the thermal noise terms enhancement at low signal-to-noise ratios due to the cubic nonlinearity. For 32-APSK a fourth power device transforms the constellation into an outer ring QPSK constellation with amplitude \( 4r_2^3 \), a middle ring composed of three points \([4r_2^3(e^{i2\pi/3}, -1, e^{-i2\pi/3})]\) and single inner point with amplitude \(-2r_1^3\). A phase shift \( \theta_P = \pi/4 \) is also required to get the right constellation orientation after the nonlinearity for the slicer correct operation. Therefore for 32-APSK the
scheme named F-4Q-DD requires $P = 4$ and $\theta_p = \pi/4$ in Fig. 6.

Figs. 8, 9, and 10 depict the simulated APSK phase error detector S-curves for $E_s/N_0 = \infty$, 15, 10 dB for 4+12-APSK 4Q-DD, 4+12-APSK C-4Q-DD, and 4+12+16-APSK F-4Q-DD schemes respectively. It should be remarked the regular S-curve sawtooth shape of the C-4Q-DD in absence of noise with an unstable lock point at $\pm$ 15 degrees corresponding to the outer ring phase ambiguity (360/12=30). The more irregular shape of the 4Q-DD phase detector of Fig. 8 can be explained by the simultaneous presence of 4+12-APSK inner and outer ring points which show different angular distance from the slicer boundaries. This does not happen for the C-4-DD PED for which inner ring points collapse in the origin. For both PEDs noise impact results in the classical smoothing of the saw tooth shape due to the decision errors. It should be remarked that in the absence of thermal noise the S-curve slope of the C-4Q-DD scheme is enhanced by a factor of about 3.8 compared to the 4Q-DD one. 32-APSK results are shown in Fig. 10: it is remarked that the narrower S-curve stable region due to the higher constellation order. Nonetheless PED sensitivity to phase error appears very good.

Fig. 11 illustrates the phase error (jitter) standard deviation obtained by simulation for 4+12-APSK and 4+12+16-APSK and the 4Q-DD, C-4Q-DD and F-4Q-DD phase estimators. These results indicate that the C-4Q-DD PED provides a remarkable phase error standard deviation improvement factor of about $3$ at $E_s/N_0 = 10$ dB compared to the 4Q-DD scheme of [24]. At $E_s/N_0 = 20$ dB the improvement corresponds to almost one order of magnitude. As we observe, a phase jitter standard deviation of about 0.4 degrees can be achieved at $E_s/N_0 = 10$ dB for 32-APSK using the F-4Q-DD PED are also remarkably good. It should be recalled that 32-APSK will typically operate at higher $E_s/N_0$ than 4+12-APSK.

In principle the proposed phase estimators can also work in feed-forward mode by performing a moving average of the complex error signal contained in the argument of the imaginary function in the phase error expression above, and then extracting the complex signal phase. However, similarly to what is seen in Figs. 8-10, the discriminator S-curve shape is dependent on the actual $E_s/N_0$. While the shape has no major impact on the feedback configuration (apart from the above mentioned loop gain changes), for the feed-forward estimator the non-linear S-curve characteristic and its slope dependencies on the SNR implies a biased estimator for non-zero phase errors. This major drawback, due to low SNR quadrant decision errors, renders impractical the exploitation of the feed-forward phase estimator scheme.

VI. NUMERICAL RESULTS

In this section we describe some numerical results that illustrate the performance of coded APSK modulations with the receiver elements described in Sect. V. As a benchmark we briefly describe some results based on concatenated Reed Solomon and Trellis Coder. When not stated otherwise we assume that roll-off factor of the SRRC filter is 0.35.

The following pre- and post-compensation configurations have been investigated: 1) “Classical” 16-QAM with linear equalization; 2) “Classical” 16-QAM with a third order 30 tap Volterra nonlinear equalizer, with LMS linear adaptation step $\alpha_{lin} = 10^{-3}$ and a nonlinear adaptation step $\alpha_{nonlin}$ determined by $10^{-6}$ (see [10], [12], [15] for details on practical implementation); 3) 4+12-APSK with linear equalization; 4) 4+12-APSK with nonlinear equalization; 5) 4+12-APSK with linear equalization and modified TC decoder metrics [14]; 6) 16-QAM with nonlinear equalization [12] and modified TC decoder metrics [14]; 7) 4+12-APSK with static pre-compensation and linear equalization; 8) 4+12-APSK with static pre-compensation and nonlinear equalization [12].

A. End-to-end Simulator Description

The system under consideration is depicted in Fig. 12. The sequence of steps in the transmitter is as follows. The binary

\[ \text{Phase error standard deviation} [\text{deg}] = \frac{E_s}{N_0} \]

\[ \text{Phase error S-curve} \]

\[ \text{Fig. 10. Simulated APSK phase error detector S-curve for} \ E_s/N_0 = \infty, 15, 10 \text{ dB; 4+12+16-APSK F-4Q-DD scheme.} \]

\[ \text{Fig. 11. APSK simulated phase jitter for: 16-APSK 4Q-DD, 16-APSK C-4Q-DD,} \]

\[ \text{32-APSK, F-4Q-DD estimators: [} B_{L0} \cdot T_s(\infty) = 10^{-4}. \]

\[ \text{Performance results for 32-APSK using the F-4Q-DD PED are also remarkably good. It should be recalled that 32-APSK will typically operate at higher} \ E_s/N_0 \text{ than 4+12-APSK.} \]

\[ \text{In principle the proposed phase estimators can also work in feed-forward mode by performing a moving average of the complex error signal contained in the argument of the imaginary function in the phase error expression above, and then extracting the complex signal phase. However, similarly to what is seen in Figs. 8-10, the discriminator S-curve shape is dependent on the actual} \ E_s/N_0. \]

\[ \text{While the shape has no major impact on the feedback configuration (apart from the above mentioned loop gain changes), for the feed-forward estimator the non-linear S-curve characteristic and its slope dependencies on the SNR implies a biased estimator for non-zero phase errors. This major drawback, due to low SNR quadrant decision errors, renders impractical the exploitation of the feed-forward phase estimator scheme.} \]
information data bits \( a_k \), at rate \( R_b \), enter a serial-to-parallel device (S/P) generating three parallel streams at rate \( R_b/3 \). An encoder with rate \( r = 3/4 \) generates four parallel binary symbol streams at rate \( R_s = R_b/3 \), that are mapped through a Gray mapping onto the APSK constellation generator. The modulator I-Q streams are at symbol rate \( R_s = R_b/3 \). 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 in Sect. II. White noise representing 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 \( N_s = T_s/T_{\text{clock}} \) samples/symbol. In practical implementations the number of samples/symbol \( N_s \) depends on the characteristic of 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 sub-sampled at two samples per symbol (on-time and early) by a digital interpolator [21] driven by the timing error detector unit. 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 performing only amplitude estimation. A dedicated phase error detector is present at the output of the digital AGC. If required, QAM frequency estimators are applicable to our case [27]. The turbo decoder represents the last demodulator block.

**B. Trellis-coded APSK**

We first consider a standard trellis-coded modulation (TCM) scheme over APSK modulations with set-partitioning or quasi-set-partitioning binary labeling rules. 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 [28], [29]. 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. For quasi-error free (QEF) performance, e.g. BER on the order of \( 10^{-10} \), we consider a concatenated scheme for which the outer code is the Reed Solomon code RS (188, 204) and the inner code is the TCM. Then, the inner decoder BER is in the order of \( 2 \cdot 10^{-4} \) [2]. The selected 16-state optimal trellis code for 16-QAM provides performance very close to that of the 64-state binary pragmatic trellis code selected for the DVB-SNG standard [2].

At the receiver side, we used the NDA approach of [20] for timing recovery, while for amplitude and phase recovery

---

**Fig. 12.** End-to-end block diagram.
we used the approach described in [14], with the following parameters: AGC adaptation step $\gamma_a = 10^{-3}$, symbol clock estimator loop noise bandwidth $B_{L_T} = 10^{-3}$ and phase estimator loop noise bandwidth $B_{L_\phi}(\infty) = 5 \cdot 10^{-4}$.

Computer simulations for the 16-state TCM are reported in Fig. 13. One should consider a required trellis decoder output BER of $2 \cdot 10^{-4}$ to find the corresponding $E_b/N_0|_{\text{sat}}$. The proposed 16-APSK constellation with pre-compensation achieves a gain of 3.1 dB over conventional 16-QAM with linear complex equalization (amplitude and phase). The trellis coded 16-APSK performs slightly better than the double-ring 16-APSK scheme proposed over the AWGN linear channel [9]. However, in the nonlinear channel, simulation results demonstrate the superior performance of the proposed 16-APSK modulation. Notice that at the BER of interest, the performance gap between the optimal pre-compensated 16-APSK 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 static simple centralized pre-compensation for 16-APSK can still provide about 0.5 dB performance improvement, decentralized nonlinear equalization at the demodulator, despite its higher complexity, yields only marginal improvement.

C. Turbo-coded APSK

We now report some results obtained by computer simulation for a turbo-coded APSK system with bit-interleaved coded modulation (BICM). 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 [18], [30]. In particular, we study the serial turbo-code devised in [9] with outer code $C^O$ the 16 states rate $r_O = 3/4$ convolutional code, and inner code $C^I$ the 2 states $r_I = 1$ recursive convolutional code (accumulator). We consider frame transmission, with 16384 information bits per frame. The corresponding spectral efficiency is $R = 3$ bit/s/Hz. We do not perform demapping iterations, and we consider Gray and quasi-Gray labelings [18], [30]. The following demodulator synchronizer settings have been adopted throughout this section: $\gamma_a = 10^{-3}$, $B_{L_T} = 10^{-3}$, $B_{L_\phi} = 5 \cdot 10^{-5}$. As for trellis coded modulation we used a third-order timing interpolator. If not specified, the number of decoding iterations is set to 10. Fig. 14 shows the performance of such turbo-coded APSK system over a nonlinear satellite channel with IBO = 2 dB. For the sake of completeness, we provide also results for the linear AWGN channel, with and without the NDA synchronizer.

We first observe that in the AWGN channel, QEF performance gain is about 1.65 dB with respect to the 16-states TCM. Notice also, that due to the concatenation with the RS code, the transmission rate corresponding to the TCM is reduced approximately by a about 8%. This amounts to a total gain of about 2.0 dB, for an increase in complexity of a factor 20-40.

Due to the marginal gain offered by nonlinear Volterra equalization for the case of TCM, we did not consider it for the turbo-coded APSK. The IBO operating point was optimized by simulation, as shown in [9], [14]. The $E_b/N_0$ was selected in the 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. (3)) $[E_b/N_0]_{\text{eff}}$. Results corresponding to the BER performance with the optimal IBO = 2 dB, are shown in Figure 14. Thanks to the combination of a powerful turbo code and a robust 16-APSK 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 $[E_b/N_0]_{\text{imp}} = 6.9$ 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 represents 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. The adoption of the more complex dynamic pre-distortion technique with memory $Q = 5$ brings an extra gain of about 0.8 dB in terms of $E_b/N_0|_{\text{sat}}$ at the optimum operating point corresponding now to $\text{IBO} = 1$ dB.

VII. SUMMARY AND CONCLUSIONS

Extensive analysis and simulations for coded APSK modulations, with particular emphasis on its applicability to satellite broadband communications have been presented in this paper. The impact of typical satellite nonlinearities and techniques to counteract their impact on the demodulator performance has been analyzed. Also simple yet effective techniques for digital demodulator synchronization recovery have been devised and their performance analyzed. It is shown that coded APSK with simple digital pre-distortion techniques based on look-up tables can achieve very good performance with satellite HPA driven at saturation (16-APSK) or with limited back-off (32-APSK). The simple digital algorithms for AGC, clock and phase recovery are shown to cause very small demodulator impairment compared to theoretical performance. The proposed turbo coded APSK scheme is shown to provide remarkable advantages in terms of power saving when used over typical satellite nonlinear channels.

For the case of carrier phase estimation, we have proposed a new class of closed-loop non-data-aided (NDA) synchronizers for APSK modulations and evaluated some of their properties (S-curves, phase jitter). We show that the new APSK phase synchronizers show no phase jitter floor due to pattern noise. We finally evaluate the overall performance of the proposed turbo-coded APSK demodulator by means of an end-to-end computer simulator, and we show an improvement of about 5 dB in power and 8 % in spectral efficiency compared to conventional Trellis-Coded 16-QAM.

REFERENCES


Riccardo De Gaudenzi was born in Italy in 1960. He received his Doctor Engineer degree (cum Laude) in electronic engineering from the University of Pisa, Italy in 1985 and the PhD from the Technical University of Delft, The Netherlands in 1999. From 1986 to 1988 he was with the European Space Agency (ESA), Stations and Communications Engineering Department, Darmstadt (Germany) where he was involved in satellite telecommunication ground systems design and testing. In particular, he followed the development of two new ESA’s satellite tracking systems. In 1988, he joined ESA’s Research and Technology Centre (ESTEC), Noordwijk, The Netherlands where in 2000 he has been appointed head of the Communication Systems Section and since 2005 he is Head of the RF Payload and Systems Division. The division is responsible for the definition and development of advanced satellite system, subsystems and technologies for telecommunications, navigation and earth observation applications. In 1996 he spent one year with Qualcomm Inc., San Diego USA, in the Globalstar LEO project system group under an ESA fellowship. His current interest is mainly related with efficient digital modulation and access techniques for fixed and mobile satellite services, synchronization topics, adaptive interference mitigation techniques and communication systems simulation techniques. From 2001 to 2005 he has been serving as Associate Editor for CDMA and Synchronization for IEEE Transactions on Communications. He is co-recipient of the VTS Jack Neubauer Best System Paper Award from the IEEE Vehicular Technology Society.

Albert Guillén i Fàbregas was born in Barcelona, Catalonia, Spain, in 1974. He received the Telecommunications Engineering Degree and the Electronics Engineering Degree from Universitat Politécnica de Catalunya, Barcelona, Catalunya, Spain, and the Politecnico di Torino, Turin, Italy, respectively, both in 1999, under the Double Degree Socrates-Erasmus Scholarship Program of the European Community, and the PhD. in Communication Systems from Ecole Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland, in 2004. From August 1998 to March 1999, he conducted his Final Research Project at the Center for Communications and Signal Processing Research (CCSPR), at the New Jersey Institute of Technology (NJIT), Newark, NJ, supported with a Nokia-NJIT Research Fellowship. He was with Telecom Italia Research Laboratories, TILAB (old CSELT), Torino, Italy, from November 1999 to June 2000. From September 2000 to May 2001, he was with the European Space Research and Technology Centre (ESTEC), of the European Space Agency (ESA), Noordwijk, The Netherlands. During his doctoral studies, from 2001 to 2004, he has been a Research and Teaching assistant at the Mobile Communications Department, Institut Eurécom, Sophia-Antipolis, France. From June 2003 to July 2004 he has been a visiting scholar at the Communications Theory Lab at EPFL. Since September 2004 he is at the Institute for Telecommunications Research, University of South Australia, Mawson Lakes, Australia, as a Research Fellow. During June-July 2005 he has held a visiting appointment at Ecole Nationale Supérieure des Télécommunications, Paris, France. His specific research interests are in the area of communication theory, information theory, coding theory, digital modulation and signal processing techniques, particularly with wireless terrestrial and satellite applications.

Dr. Guillén i Fàbregas received a pre-doctoral Research Fellowship of the Spanish Ministry of Education to join ESTEC-ESA. He received the Young Authors Award of the 2004 European Signal Processing Conference EUSIPCO 2004, Vienna, Austria and the 2004 Nokia Best Doctoral Thesis Award in Mobile Internet and 3rd Generation Mobile Solutions from the Spanish Institution of Telecommunications Engineers. He is also a member of the ARC Communications Research Network (ACoRN).

Alfonso Martinez was born in Zaragoza, Spain, in October 1973. He graduated in Electrical Engineering from the University of Zaragoza, Spain, in 1997. From 1998 to 2002 he worked at the research center of the European Space Agency in Noordwijk, the Netherlands. He made research on coded modulation for satellite systems. Since 2003 he is a PhD student at Technische Universiteit Eindhoven, the Netherlands, where he is currently working in optical communication theory.