
Any correspondence concerning this service should be sent to the repository administrator: staff-oatao@inp-toulouse.fr
Trade-off between spectrum efficiency and link unavailability for hierarchical modulation in DVB-S2 systems

Hugo Méric∗†, Jérôme Lacan†∗, Caroline Amiot-Bazile§, Fabrice Arnal‡ and Marie-Laure Boucheret†∗

∗TeSA, Toulouse, France
†Université de Toulouse, Toulouse, France
‡Thales Alenia Space, Toulouse, France
§CNES, Toulouse, France

Email: hugo.meric@isae.fr, jerome.lacan@isae.fr, caroline.amiot-bazile@cnes.fr, fabrice.arnal@thalesaleniaspace.com, marie-laure.boucheret@enseeiht.fr

Abstract—Broadcasting systems have to deal with channel variability in order to offer the best spectrum efficiency to the receivers. However, the transmission parameters that optimize the spectrum efficiency generally leads to a large link unavailability. In this paper, we study the performance of hierarchical and non-hierarchical modulations in terms of spectrum efficiency and link unavailability for DVB-S2 systems. Our first contribution is the design of the hierarchical 16-APSK for the DVB-S2 standard. Then we introduce the link unavailability to compare the performance of hierarchical and non-hierarchical modulations in terms of spectrum efficiency and link unavailability. The results show that hierarchical modulation is a good alternative to non-hierarchical modulation for the DVB-S2 standard.

I. INTRODUCTION

In most broadcast systems, all the receivers do not experience the same signal-to-noise ratio (SNR). For instance, in satellite communications the channel quality decreases with the presence of clouds in Ku or Ka band, or with shadowing effects of the environment in lower bands. The first solution for broadcasting is to design the system for the worst-case reception. However, this solution does not take into account the variability of channel quality. This leads to poor performance as the receivers with good reception do not exploit their full potential. Two other schemes have been proposed in [1]: time division multiplexing with variable coding and modulation, and superposition coding. Time division multiplexing, or time sharing, allocates to each user a fraction of time where it can use the channel with any modulation and error protection level. In [1], Cover introduced superposition coding in order to improve the previous scheme. When communicating with two receivers, the principle is to superimpose information for the receiver with the best SNR. This superposition can be done at the forward error correction level or at the modulation level by sharing the available energy among several data streams which are sent simultaneously in the same band. Hierarchical modulation is a practical implementation of superposition coding at the modulation level. The principle is illustrated in Figure 1 with a non-uniform 16-QAM symbol. Note that hierarchical modulation often relies on non-uniform constellations where the symbols are not uniformly distributed in the space.

Although superposition coding was introduced to improve capacity, hierarchical modulation is mainly used for several other applications. Firstly, it is often used to provide unequal protection. The idea is to allocate unequal amounts of energy between the transmitted streams. The more power is given to a stream, the easier it is decoded. If two streams are merged at the modulation level, the stream with more energy is called the high priority (HP) stream and is dedicated to receivers with poor channel quality. The other stream is called the low priority (LP) stream and requires a larger SNR to be decoded. For instance in Figure 1, the HP stream is used to select the quadrant and the LP stream selects the position inside the quadrant. An example of practical application to provide unequal protection is given in [2], where H.264/SVC encoded video [3] is protected using hierarchical modulation. The base layer of the video is transmitted in the HP stream, while the enhanced layer is carried by the LP stream. This approach allows each receiver to decode a video quality commensurate with its channel quality. Another application of hierarchical modulation is backward compatibility [4], [5]. The DVB-S2 standard [6] is called upon to replace DVB-S, but many DVB-S receivers are already installed. The hierarchical modulation helps to the migration by allowing the DVB-S receivers to operate. Finally, hierarchical modulation has several other applications: providing local content [7], performance im-
The DVB-S2 standard mainly relies on variable coding and modulation which is an implementation of time sharing [6]. This functionality combines LDPC codes with a variety of non-hierarchical modulation formats: QPSK, 8-PSK, 16-APSK and 32-APSK. As already mentioned, the standard also considers hierarchical modulation (with the hierarchical 8-PSK) but only for backward compatibility purpose.

This paper focuses on the trade-off between spectrum efficiency and link unavailability for hierarchical and non-hierarchical modulations in a DVB-S2 system. Indeed, the transmission parameters that offer the best spectrum efficiency generally require a large SNR to be decoded which leads to a large link unavailability. In Section II, we propose the hierarchical 16-APSK to improve the performance of the DVB-S2 standard. In Section III, we introduce the link unavailability to compare the performance of hierarchical and non-hierarchical modulations in terms of spectrum efficiency and link unavailability. Finally, Section IV concludes the paper by summarizing the results.

II. HIERARCHICAL 16-APSK

A. Introduction

The DVB-S2 standard considers the hierarchical 8-PSK for backward compatibility purpose. However, this modulation does not provide any spectrum efficiency improvement in comparison to the other modulations in the standard. The DVB-S2 standard also considers the 16-APSK modulation that was preferred to the 16-QAM modulation as the 16-APSK has better performance on a non-linear transponder and comparable performance on linear channel [6]. As the 16-APSK is already defined in the standard, we propose to introduce the hierarchical 16-APSK illustrated in Figure 2. The constellation parameters are the ratio between the radius of the inner and outer ring and the symbol energy is given by the energy of a QPSK modulation, where the constellation points are located at the barycenter of the four points in each quadrant. We now compute the positions of these barycenters. Using the polar coordinates, the barycenter in the upper right quadrant is

\[ z_b = e^{i\pi/4} \frac{R_1 + R_2 + 2R_2 \cos(\theta)}{4} \]  

(2)

We search to introduce the symbol energy \( E_s \) in (2). For the 16-APSK, the symbol energy is expressed as

\[ E_s = \frac{4R_1^2 + 12R_2^2}{16} = \frac{1 + 3\gamma^2}{4} R_1^2. \]  

(3)

Then combining (2) and (3), the distance of the barycenter to the origin is

\[ d_B = |z_b| = \frac{1 + \gamma(1 + 2 \cos(\theta))}{4} 2\sqrt{E_s}. \]  

(4)

Thus, the energy of the HP stream is given by

\[ E_{hp} = d_B^2 = \frac{(1 + \gamma(1 + 2 \cos(\theta)))^2}{4(1 + 3\gamma^2)} E_s. \]  

(5)

In Equation (6), we introduce \( \rho_{hp} \) the ratio between the energy of the HP stream \( E_{hp} \) and the symbol energy \( E_s \). The equation between the energy allocated to the HP stream and the constellation parameters is

\[ \rho_{hp} = \frac{E_{hp}}{E_s} = \frac{(1 + \gamma(1 + 2 \cos(\theta)))^2}{4(1 + 3\gamma^2)}. \]  

(6)

As the HP stream contains more energy than the LP stream, we verify that \( \rho_{hp} \geq 0.5 \). For a given \( \rho_{hp} \), we search the \((\gamma, \theta)\) pairs \((\gamma \geq 1 \text{ and } \theta \geq 0)\) solution of (6). In order to solve (6), we transform the equation as follow

\[ \cos \theta = \frac{1}{2} \left( \sqrt{4\rho_{hp}(1 + 3\gamma^2)} - 1 \right) \]  

(7)
The term \( \cos \theta \) is a function that depends on \( \gamma \) and \( \rho_{hp} \). We note \( f(\gamma, \rho_{hp}) \) this function.

C. Resolution of the energy equation

We are now interested to determine the set of \((\gamma, \theta)\) pairs solution of (7) where \( \rho_{hp} \) is known. The principle is to express \( \theta \) as a function of \( \gamma \).

We now search when the condition \(-1 \leq f(\gamma, \rho_{hp}) \leq 1\) is verified in order to use the \( \arccos \) function. The derivative of \( f \) shows that the function \( f(\gamma, \rho_{hp}) \) is an increasing function of \( \gamma \) when \( \rho_{hp} \) is set. Using the facts that \( \gamma = R_2/R_1 \geq 1 \) and \( \rho_{hp} \geq 1/2 \), we can write

\[
f(\gamma, \rho_{hp}) \geq f(1, \rho_{hp}) = \frac{1}{2} (4\sqrt{\rho_{hp}} - 2) \geq \sqrt{2} - 1 \quad (8)
\]

Thus the function \( f(\gamma, \rho_{hp}) \) always verifies \(-1 \leq f(\gamma, \rho_{hp}) \leq 1\). We now study an upper bound of \( f \). First of all, we have the following relation

\[
f(\gamma, \rho_{hp}) = \frac{1}{2} (2\sqrt{3\rho_{hp}} - 1) \quad (9)
\]

The right term is an increasing function in \( \rho_{hp} \) and equals 1 for \( \rho_{hp} = 0.75 \). Thus, for all \( \rho_{hp} \leq 0.75 \), the condition \(-1 \leq f(\gamma, \rho_{hp}) \leq 1\) is verified and the \( \arccos \) function can be used in (7). The solution of (7) for \( \rho_{hp} \leq 0.75 \) is

\[
S_{\rho_{hp}} = \{(\gamma, \arccos (f(\gamma, \rho_{hp})) | | \gamma | \geq 1\} \quad (10)
\]

When \( \rho_{hp} > 0.75 \), \( \gamma \) must stay bounded in order to verify \( f(\gamma, \rho_{hp}) \leq 1 \). To determine the limit value \( \gamma_{lim} \), we have to solve the equation

\[
f(\gamma, \rho_{hp}) = 1 \iff \frac{1}{2} \left( \frac{4\rho_{hp}(1 + 3\gamma^2) - 1}{\gamma} - 1 \right) = 1
\]

\[
\iff (12\rho_{hp} - 9)\gamma^2 - 6\gamma + (4\rho_{hp} - 1) = 0.
\]

Equation (11) is a quadratic equation with discriminant \( \Delta = 192\rho_{hp}(1 - \rho_{hp}) \). The solutions are

\[
s_{1,2} = \frac{6 \pm \sqrt{192\rho_{hp}(1 - \rho_{hp})}}{2(12\rho_{hp} - 9)} \quad (12)
\]

We keep the positive solution,

\[
\gamma_{lim} = \frac{3 + \sqrt{3\rho_{hp}(1 - \rho_{hp})}}{4(3\rho_{hp} - 3)} \quad (13)
\]

Finally, the solution of (7) for \( \gamma > 0.75 \) is

\[
S_{\rho_{hp}} = \{(\gamma, \arccos (f(\gamma, \rho_{hp})) | | \gamma \leq \gamma_{lim}\} \quad (14)
\]

Figure 3 presents two examples of \( S_{\rho_{hp}} \) with different values of \( \rho_{hp} \). When \( \rho_{hp} \) increases, the symbols in one quadrant tend to come closer. For instance, when \( \gamma = 1 \), we find that \( \theta = 38^\circ \) for \( \rho_{hp} = 0.8 \) and \( \theta = 26^\circ \) for \( \rho_{hp} = 0.9 \). Thus the symbols are closer in the case \( \rho_{hp} = 0.9 \). This implies that the HP stream is easier to decode, but on the other hand the LP stream requires a good reception to be decoded.

D. Performance of the hierarchical 16-APSK

In practical systems, several values of \( \rho_{hp} \) have to be chosen. Moreover, once the value of \( \rho_{hp} \) is known, there remains to pick one \((\gamma, \theta)\) pair in the \( S_{\rho_{hp}} \) set. In this paper, we choose the pair that minimises the decoding threshold of the HP stream averaged over all the DVB-S2 coding rates. To obtain a fast evaluation of the decoding thresholds in function of a constellation, we use the method described in [10]. Table I presents the adopted values.

\[
\begin{array}{cccc}
\rho_{hp} & 0.75 & 0.8 & 0.85 & 0.9 \\
\gamma & 2.8 & 2.3 & 1.9 & 1.6 \\
\theta & 31.5 & 28.4 & 25.1 & 20.9 \\
\end{array}
\]

Finally, the performance in terms of bit error rate (BER) of the hierarchical 16-APSK is evaluated with simulations. We use the Coded Modulation Library [11] that already implements the DVB-S2 LDPC (without the concatenated BCH outer code). The LDPC codewords are 64 800 bits long (normal FEC frame) and the iterative decoding stops after 50 iterations if no valid codeword has been decoded. Moreover, in our simulations, we wait until 10 decoding failures before computing the BER. If the BER is less than \( 10^{-4} \), then we stop the simulation. Our stopping criterion is less restrictive than in [6] (i.e, a packet error rate of \( 10^{-7} \)) because simulations are time consuming. However, our simulations are sufficient to detect the waterfall region of the LDPC and then the performance of the code. Figure 4 presents the performance of the HP and LP streams for \( \rho_{hp} = 0.8 \).

III. SPECTRUM EFFICIENCY VS LINK UNAVAILABILITY

This section addresses the trade-off between spectrum efficiency and link unavailability for DVB-S2 systems.

A. Definition of link unavailability

In this section, we seek to take into account the channel variability of a broadcast system. The (link) unavailability is in that case relevant to complete the spectrum efficiency in the choice of the transmission parameters (modulation and code rate). The unavailability is defined as the percentage of the population which can not decode any stream. Its computation requires SNR distributions of the receivers. This
defined as average spectrum efficiency and a reasonable unavailability. We consider here an transmission parameters maximising the spectrum efficiency may also be decoded by a small fraction of the population. A compromise has to be found between a good spectrum efficiency and a reasonable unavailability. We consider here an average spectrum efficiency over the population who decodes at least the HP stream. The average spectrum efficiency is defined as

\[
\text{Average Spectrum Efficiency} = \frac{\mu_{hp} \tau_{hp} + \mu_{lp} \tau_{lp}}{\tau_{hp}},
\]

(15)

where \( \mu_x \) represents the spectrum efficiency for the stream \( x \) and \( \tau_x \) is the percentage of the population decoding the stream \( x \). We assume that the transmission parameters ensure that \( \tau_{lp} \leq \tau_{hp} \). In the best case, the whole population decodes both streams so \( \tau_{hp} = \tau_{lp} = 1 \) and the average spectrum efficiency equals \( \mu_{hp} + \mu_{lp} \).

**B. Application to DVB-S2**

1) Channel model: We present a model to estimate the SNR distribution of the receivers in the Ka band. We consider the set of receivers located in a given spot beam of a geostationary satellite broadcasting in the Ka band. The model takes into account two main sources of attenuation: the relative location of the terminal with respect to the center of (beam) coverage and the weather. Concerning the attenuation due to the location, the principle is to set the SNR at the center of the spot beam (\( SNR_{max} \)) and to use the radiation pattern of a parabolic antenna to model the attenuation. An approximation of the radiation pattern is

\[
G(\eta) = G_{max} \times \left( \frac{J_1 \left( \frac{\pi \eta D}{\lambda} \right)}{\sin(\eta) \frac{\pi D}{\lambda}} \right)^2,
\]

(16)

where \( G_{max} \) is the maximum gain, \( J_1 \) is the first order Bessel function, \( D \) is the antenna diameter, \( \lambda = c/f \) is the wavelength and \( \eta \) is the angle [12]. In our simulations, we use \( D = 1.5 \) m and \( f = 20 \) GHz. Moreover, we consider a typical multispot system where the edge of each spot beam is 4 dB below the center of coverage. Assuming a uniform repartition of the population, the proportion of receivers experiencing an attenuation between two given values is computed as follows: compute the two angles, \( \eta_1 \) and \( \eta_2 \), corresponding to the two attenuation values using (16). This defines a ring as shown in Figure 5. The proportion is finally given by the ratio of the ring area over the spot beam area.

Figure 6, provided by the CNES (the french space agency), shows the attenuation distribution in the broadcasting satellite service band. More precisely, it is a temporal distribution for a given location in Toulouse, France. In our work, we assume that the SNR distribution for the receivers in the beam coverage at a given time is equivalent to the temporal distribution at a given location.

Finally, our model combines the two sources of attenuation previously described, location and weather, to estimate the SNR distribution. From a set of receivers, we first compute the attenuation due to the location. Then, for each receiver we draw the attenuation caused by the weather according to the distribution in Figure 6.
2) Transmission parameters: In our simulations, we use all the coding rates of the DVB-S2 standard and the following modulations: QPSK, 16-APSK and hierarchical 16-APSK.

3) Results: Figure 7 presents the performance of the hierarchical 16-APSK in terms of unavailability and average spectrum efficiency. Each curve has been obtained with 50,000 receivers where the SNR distribution is drawn according to the previous model. For each figure, we set the SNR at the center of the spot beam with clear sky condition, \( SNR_{max} \).

Concerning the hierarchical modulation, once the coding rates of both streams have been chosen, we only represent the points that verify the constraint \( E_b/N_0 \text{lp} \geq E_b/N_0 \text{hp} \), where \( E_b/N_0 \text{lp} \) is the decoding threshold of the stream \( x \). Thus the unavailability only depends on the coding rate of the HP stream and the constellation parameters. This explains the shape of the curves in Figure 7. Note that the coding rate of the LP stream has only an impact on the average spectrum efficiency.

For a given HP stream coding rate, when \( \rho_{lp} \) increases, the decoding threshold of the HP stream decreases as well as the link unavailability. Thus the minimum unavailability is obtained for \( \rho_{lp} = 0.9 \).

It is also important to point out that, when the coding rate of the LP stream increases (for a given HP stream coding rate), the average spectrum efficiency does not necessarily increase. Indeed, when the LP stream coding rate increases, the term \( \mu_{lp} \) in (15) increases, however the term \( \tau_{lp} \) decreases.

Finally, the 16-APSK modulation obtains the worst results. This can be explained as the DVB-S2 standard only considers this modulation with coding rates greater or equal to 2/3 [6]. Thus a good reception is needed to decode the 16-APSK modulation. The performance comparison between the QPSK and the hierarchical 16-APSK modulations depends on the value of \( SNR_{max} \). For \( SNR_{max} = 5 \) dB which leads to very low SNR for all the receivers, the QPSK modulation obtains the best results (see Figure 7a). Even if the hierarchical modulation competes at the unavailability level, it does not improve the spectrum efficiency. In fact, for low SNR values, the LP stream can only be decoded with small coding rates which does not improve the performance. However, the results are in favor of the hierarchical 16-APSK for \( SNR_{max} = 10 \) dB. For instance, if we consider an unavailability of 0.02% (classical unavailability targets of satellite broadcasting systems are below 1%), the hierarchical modulation almost doubles the average spectrum efficiency. The performance improvement is due to the LP stream that can be decoded with larger coding rate than in the \( SNR_{max} = 5 \) dB configuration.

IV. CONCLUSION

In this paper, we compare the performance of non-hierarchical and hierarchical modulations in terms of spectrum efficiency and link unavailability. First, we present the hierarchical 16-APSK to improve the performance of the DVB-S2 standard. By considering the energy allocated to the HP stream, we show how to compute the constellation parameters. Then, we introduce the link unavailability using SNR distributions. Using this notion, we compare different transmission parameters with two criteria: average spectrum efficiency and unavailability. Our results point out that hierarchical modulation may provide better performance than classical non-hierarchical. In a future work, we expect to investigate the impact of the satellite channel non-linearity on the performance of the hierarchical modulation.

REFERENCES