Multi-Resolution Modulation: An Optimization Criterion
Based on Information Theory

Marie Zwingelstein-Colin, Mohamed Gharbi, Marc Gazalet
IEMN/DOAE – UMR CNRS 8520
Université de Valenciennes
Le Mont Houy
59313 Valenciennes Cedex, France
Marie.Zwingelstein-Colin@univ-valenciennes.fr

Abstract - Although the degree of protection for transmitted bits offered by Multi-Resolution Modulation (MRM) depends on the bits' perceptual importance, the channel capacity concept introduced by Shannon intentionally excludes the notion of information importance. In response to this exclusion, the authors introduce an alternate information theory-based quantity, called "weighted channel capacity", which explicitly integrates the notion of information importance and which can be used as an MRM optimization criterion. To accomplish this, the chain rule of mutual information was used to weight the influence of each bit according to its perceptual importance. Numerical results for the "weighted channel capacity" quantity are provided for the Hierarchical Pulse Amplitude Modulation (H-PAM) class on the AWGN channel.

Index terms - Multi-resolution modulation, information importance, channel capacity, hierarchical modulation.

I. INTRODUCTION

Multi-Resolution Modulation (MRM), and Unequal Error Protection (UEP) systems in general, seeks to provide different levels of protection for the transmitted data according to the data's perceptual importance, as determined by the receiver, with the overall goal of improving the receiver's satisfaction level [1-4]. In High Definition TV (HDTV) hierarchical transmission, for instance, the joint use of MRM and hierarchical video coding allows a progressive degradation of the video reception quality when reception conditions deteriorate rather than an abrupt degradation (HDTV quality in good reception conditions, TV or even Personal Video (PV) quality in bad reception conditions) [5,6].

In [5], the authors underlined the need for a performance criterion for which the parameters of a multi-resolution broadcast transmission system could be optimized. They stated that finding such a criterion would be problematic, and pointed out the lack of a clear cost function. This paper offers a partial solution to the problem raised in their article. It proposes a new theoretical quantity that resolves the incompatibility between MRM – in which the idea of the perceptual importance of information is implicit – and Shannon's channel capacity concept [7], which intentionally excludes the notion of information importance. This new quantity is called "weighted channel capacity" and can be used as an MRM optimization criterion.

In the following section, the idea of the "weighted channel capacity" quantity as an MRM optimization criterion is developed. In section III, the criterion's numerical results are presented for the Hierarchical Pulse Amplitude Modulation (H-PAM) class on the AWGN channel.

II. MOTIVATION AND DEFINITION OF THE CONCEPT

Though not always mentioned explicitly in the literature, MRM implies a systematic understanding that the different transmitted data do not have the same perceptual importance for the receiver. Given that transmitted data contain information (as defined by information theory), integrating the idea of "relative importance" into the concept of "information" seems logical. This insight eventually led the authors to the concept of "weighted channel capacity".

A. System model:

In the following sections, random variables are denoted by upper case letters and their realizations, by lower case letters.

This paper focuses on the Additive White Gaussian Noise (AWGN) channel with discrete-valued input signal points \( a \in \mathbb{R}^D \) issued from an M-ary MRM modulation scheme in a \( D \)-dimensional signal space, and continuous outputs \( y \). The channel output signal points \( y \) come from the alphabet \( Y = \mathbb{R}^D \) of real numbers in \( D \) dimensions. A bijective mapping of the input binary vectors \( x = (x_0, x_1, \ldots, x_{l-1}) \), \( x_i \in \{0,1\} \) and \( M = 2^{l-1}, l>1 \) to input signal points \( a \) must be defined. Obviously, multi-resolution is achieved when the input signal points \( a \) are not uniformly distributed. Fig. 1 shows an example of the input signal points alphabet (constellation) and the mapping of a generalized hierarchical 8-PAM (H-8PAM) modulation scheme.
\( D = 1 \) with Karnaugh map style Gray coding [8]. The parameters that affect the multi-resolution aspect (hierarchical parameters) can be defined as \( \lambda_i = d_i / d_0 \) and \( \lambda_2 = d_2 / d_0 \), \( 0 \leq \lambda_2 \leq \lambda_1 \leq 1 \), which means that the most important bit \( x^0 \) is less subject to error than the bit \( x^1 \), which is itself better protected than the least important bit \( x^2 \).

It is assumed that a multi-resolution source coding scheme is used to represent the source in a hierarchy of resolutions, in which each bit \( x^l \) belongs to one resolution level. If \( x^0 \) is the most important bit, the perceptual importance of bit \( x^l \) in relation to bit \( x^0 \) is defined as \( \alpha_i \), with \( 0 \leq \alpha_i \leq \alpha_1 \leq 1 \), \( 1 \leq l \leq L \). (In the example given in Fig. 1, \( I \) is equal to 3). Notice that the values of the \( \alpha \) parameters depend exclusively on the source coding scheme. The problem of their evaluation is not the object of the work presented in this paper.

### B. Weighted channel capacity

In his fundamental paper [7], Shannon introduced the useful definition of "channel capacity": the upper bound for the rate at which information should be able to be transmitted reliably through a given communication channel. Shannon's definition, based exclusively on the statistical structure of the channel and the message to be transmitted, intentionally excludes the idea of information meaning, and thus the perceptual importance of the information for the final receiver. Yet this notion is the basic motivation for developing joint source-channel coding schemes such as hierarchical video transmission systems, which use MRM extensively. A new quantity needs to be defined for such systems, one that, though inspired by Shannon's channel capacity concept, also adequately integrates the notion of the relative perceptual importance of the data (\( \alpha_i \) parameters). This paper responds to this need by proposing the "weighted channel capacity" \( C_w \).

Shannon's channel capacity can be written:

\[
C = \max \left( I(X;Y) \right) \text{ bit/symbol, where } P_x(x) \text{ denotes the probability of the source. Using the chain rule of mutual information [9, p.22], the mutual information } I(X;Y) \text{ can be written:}
\]

\[
I(X;Y) = I(X^0;Y) + I(X^1;Y | X^0) + \cdots + I(X^L;Y | X^0 \cdots X^{L-1}) \tag{1}
\]

As stated in [10], one interpretation of (1) leads to the concept of equivalent parallel channels, where each equivalent channel supports the transmission of the individual bit \( x^l \), provided that bits \( x^0 \ldots x^{l-1} \) are known. Figure 2 illustrates this concept for the generalized H-8PAM shown in figure 1. The actual channel is represented in Fig 2.a). The equivalent channel for bit \( x^2 \) (Fig 2.b) is one of four asymmetric 2-PAM mappers (depending on the values of bits \( x^0 \) and \( x^1 \)) followed by the actual noisy channel. The equivalent channel for bit \( x^1 \), bit \( x^2 \) acts as a disturber. Consequently, the equivalent channel for bit \( x^1 \) (Fig 2.c) can be seen as one of two asymmetric 2-PAM mappers (depending on the value of bit \( x^0 \) plus an additive discrete noise source \( d_n \), representing the noisy effect of bit \( x^1 \), followed by the actual noisy channel. The equivalent channel for bit \( x^0 \) (Fig 2.d) follows the same approach, where the additive discrete noise \( d_{n0} \) represents the combined noisy effects of bits \( x^2 \) and \( x^1 \).

This view of equivalent channels is essential for introducing the concept of "weighted channel capacity" because its indicates explicitly the proportion of each bit \( x^l \) issued from the hierarchical source coder in the Shannon’s channel capacity. Consequently, the relative importance (parameter \( \alpha_i \) of each bit \( x^l \) can be taken into account very simply by weighting the average mutual information of the associated equivalent channel by the factor \( \alpha_i \). This weighting operation yields the following definitions of "weighted mutual information" \( I_w \) and "weighted channel capacity" \( C_w \):

\[
I_w(X;Y) = I(X^0;Y) + \alpha_i I(X^1;Y | X^0) + \cdots + \alpha_{L-1} I(X^{L-1};Y | X^0 \cdots X^{L-2}) \tag{2}
\]

\[
C_w = \max \left( I_w(X;Y) \right) \text{ bit/symbol.} \tag{3}
\]

In these definitions, \( I_w = I \) and \( C_w = C \) when all the bits have an equal relative importance, i.e. when \( \alpha_i = 1 \forall i \). In all other cases, \( C_w < C \).

### III. Application for Optimizing H-PAM Modulations on the AWGN Channel

This section describes the computation of the "weighted channel capacity" for the H-PAM multi-resolution modulation and shows how the new quantity can be used as a criterion for
optimizing the hierarchical parameters \( \lambda_i \). The inputs for this optimization problem are the SNR and the relative importance factors \( \alpha_i, 1 \leq i \leq 7 \).

Assuming that only bits of equiprobable occurrence are of interest (which is the case for practical systems), the weighted channel capacity \( C_w \) can be computed easily from (2). This computation yields

\[
I(X'; Y | X^0, X^1 \cdots X^7) = H(Y | X^0, X^1 \cdots X^7) - H(Y | X^0, X^1 \cdots X^7 | Y)
\]

where the differential entropy \( H \) can be expressed as

\[
H(Y | X^0, X^1 \cdots X^7) = -\int f_{X^0 \cdots X^7} (y | X^0 \cdots X^7) \log \left( f_{X^0 \cdots X^7} (y | X^0 \cdots X^7) \right) dy
\]

where \( f_{X^0 \cdots X^7} (y | X^0 \cdots X^7) \) is the transition probability density function of any one of the equivalent channels for bit

\( x^{i+1} \).

For simplicity, the "weighted channel capacity" computations for the two least-complicated H-PAM modulation cases are shown here:

- in H-4PAM, two bits, \( x^0 \) and \( x^1 \), are mapped into the four non-equally distributed amplitude levels shown in Fig.3; the hierarchical parameter is defined as \( \lambda = d_1/d_0 \); and \( \alpha_1 = \alpha \) is the importance of bit \( x^1 \) relative to bit \( x^0 \).

- in H-8PAM, three bits—\( x^0, x^1 \), and \( x^2 \)—are mapped into the eight amplitude levels shown in Fig.1. This case has two sub-cases:
  - H-8PAM with a 2-level hierarchy \( \Rightarrow \alpha_1 = \alpha_2 = \alpha \). In this case, the source is hierarchically coded into two bitstreams of unequal importance, with the most important bitstream composed of bits \( x^0 \) and the less important bitstream composed of bits \( x^1 \) and \( x^2 \). Since \( x^1 \) and \( x^2 \) have the same weight, \( \lambda_1 = \lambda_2 = \lambda \). In this case, the source is hierarchically coded into two bitstreams of unequal importance, with the most important bitstream composed of bits \( x^0 \) and the less important bitstream composed of bits \( x^1 \) and \( x^2 \). Since \( x^1 \) and \( x^2 \) have the same weight, \( \lambda_1 = \lambda_2 = \lambda \).

A. H-4PAM

For a given SNR and a given relative importance level \( \alpha \), Fig.4 represents the evolution of the weighted channel capacity for the H-4PAM case \( C_w(\alpha, \lambda, \text{SNR}) \) with the hierarchical parameter \( \lambda \) and for given importance level \( \alpha \) and SNR. Each of these curves reaches a maximum for the optimal value

\( \lambda = \lambda_{\text{opt}} \) of the hierarchical parameter. Fig.5, which is of practical interest, gives the optimal values \( \lambda_{\text{opt}}(\alpha, \text{SNR}) \) for the hierarchical parameter with respect to the SNR and for a given importance level \( \alpha \). Fig. 5 indicates that \( \lambda_{\text{opt}} = 0 \) when the channel is very bad and/or when the secondary data’s relative perceptual importance is very weak, which corresponds to a 2-PAM mono resolution transmission of the most important data only. On the other hand, when the channel noise level is low and/or when the secondary data’s relative perceptual importance level is high, the optimal solution tends to transmit all the data in a non hierarchical 4-PAM modulation scheme. Somewhere between these two extremes—i.e., when \( 0 < \lambda_{\text{opt}} < 1 \)—the H-4PAM modulation scheme is better than either the 2-PAM or the 4-PAM mono resolution modulation schemes. In order to measure exactly how much better, the 2-level H-PAM hierarchical gains are defined relative to both the 2-PAM and the 4-PAM, and are denoted respectively as \( G_{2,0} \) and \( G_{2,1} \), where:

\[
G_{2,i} (\alpha, \text{SNR}) = \frac{C_w(\alpha, \lambda_{\text{opt}}(\alpha, \text{SNR}), \text{SNR}) - 1}{C_w(\alpha, \lambda, \text{SNR}) - 1}, i \in \{0, 1\}
\]

The significant value of the hierarchical gain is \( G_2 = \min_i (G_{2,i}) \). An exhaustive simulation was completed for \( 0 \leq \alpha \leq 1 \) and \( 0 \leq \text{SNR} \leq 22 \text{dB} \), with the results showing that the hierarchical gain \( G_2 \) has a maximum value of 5.2% (obtained when \( \alpha = 0.2 \) and \( \text{SNR} = 9.8 \text{dB} \)). These results are presented in Fig. 6 for \( \alpha \in \{0.1, 0.2, 0.4\} \). A maximum in \( G_{2,1} \) can be noticed at very low SNR values. Although these SNR values are not practical ones, it could be valuable to understand the origin of this maximum.

B. H-8PAM with a 2-level hierarchy

For H-8PAM with a 2-level hierarchy modulation scheme, the expression for \( G_{2,i} \), given in (7), corresponds to the 2-level H-8PAM hierarchical gains relative to both the 8-PAM and the 2-PAM for \( i = 1 \) and \( i = 0 \) respectively.

Fig. 7, Fig. 8, and Fig. 9 are the respective equivalents of Fig. 4, Fig. 5, and Fig. 6 for the H-8PAM with a 2-level hierarchy. The maximum in Fig.7 is more pronounced than that of Fig.4, which seems natural since the proportion of less important data has increased significantly. For the same reason, the simulation of \( G_{2,i} \) for \( 0 \leq \alpha \leq 1 \) and \( 0 \leq \text{SNR} \leq 22 \text{dB} \) shows that the maximum for \( G_2 \) is now 8.0% (obtained when \( \alpha = 0.2 \) and \( \text{SNR} = 11 \text{dB} \)). Fig. 9 represents \( G_{2,i} \) as a function of the SNR for

\[
\begin{align*}
\text{SNR} &= 11 \text{dB} \\
\alpha &= 0.2 \\
\lambda &= \lambda_{\text{opt}}(\alpha, \text{SNR})
\end{align*}
\]

Fig. 3. Hierarchical 4-PAM constellation.
Hierarchical parameter, $\lambda$

Weighted channel capacity, $C_w$ (bit/symbol)

SNR=7 dB, $\alpha=0.4$

SNR=10 dB, $\alpha=0.2$

SNR=13 dB, $\alpha=0.1$

Fig. 4. Evolution of the weighted H-4 PAM channel capacity $C_w$ on the AWGN channel in terms of the hierarchical parameter $\lambda$.

Optimal hierarchical parameter, $\lambda_{opt}$

Fig. 5. Optimal value $\lambda_{opt}$ of the H-4 PAM hierarchical parameter in terms of the SNR.

Hierarchical gain, $G_{2,i}$ (%)

i = 0

$\alpha=0.1$

$\alpha=0.2$

$\alpha=0.4$

Fig. 6. H-4PAM hierarchical gains in relation to the 2-PAM ($i=0$) and the 4-PAM ($i=1$).

Fig. 7. Evolution of the weighted 2-level H-8 PAM channel capacity $C_w$ on the AWGN channel in terms of the hierarchical parameter $\lambda$.

Optimal hierarchical parameter, $\lambda_{opt}$

Fig. 8. Optimal value $\lambda_{opt}$ of the 2-level H-8 PAM hierarchical parameter in terms of the SNR.

Hierarchical gain, $G_{2,i}$ (%)

i = 0

$\alpha=0.1$

$\alpha=0.2$

$\alpha=0.4$

Fig. 9. 2-level H-8PAM hierarchical gains in relation to the 2-PAM ($i=0$) and the 8-PAM ($i=1$).
\(\alpha \in \{0.1, 0.2, 0.4\}\). Fig. 8 shows that the values of \(\lambda_{\text{opt}}\) for the 2-level H-8PAM are systematically lower than the corresponding H-4PAM \(\lambda_{\text{opt}}\) values, which can be interpreted to mean that it is preferable to maintain the transmission quality of the most important data (to the detriment of the less important data) if the quantity of the less important data increases significantly.

C. 3-levels H-8PAM

This section considers H-8PAM with a 3-level hierarchy multi-resolution modulation scheme. As in the two previous cases, the desired outcome is an optimization of the hierarchical parameters \(\lambda_1\) and \(\lambda_2\) (defined in figure 1) for a given value of the triplet \((\alpha_1, \alpha_2, \text{SNR})\).

Fig. 10 represents the evolution of \(\lambda_{1,\text{opt}}\) and \(\lambda_{2,\text{opt}}\) in terms of the SNR for \(\alpha = 0.5\) and for three values of \(\alpha_2\). For 0≤SNR≤22 dB, three regions can be distinguished:

- For very low SNRs (\(\lambda_{1,\text{opt}}=\lambda_{2,\text{opt}}=0\)), the “weighted channel capacity” \(C_w\) is maximized by discarding bits \(x^1\) and \(x^2\) and transmitting bit \(x^0\) in a 2-PAM scheme.
- For middle range SNRs (0<\(\lambda_{1,\text{opt}}=1\) and \(\lambda_{2,\text{opt}}=0\)), \(C_w\) is optimized by discarding bit \(x^2\) and transmitting bits \(x^0\) and \(x^1\) in a 2-level H-4PAM scheme.
- For higher SNRs (0<\(\lambda_{1,\text{opt}}<\lambda_{2,\text{opt}}<1\)), \(C_w\) is optimized by exploiting the 3-level hierarchy of the H-8PAM scheme.

As in section III.A and III.B, in order to evaluate the worth of the H-8PAM in relation to the 2-PAM, 4-PAM and 8-PAM, the 3-level H-PAM hierarchical gains are defined relative to the 8-PAM, 4-PAM and 2-PAM and are denoted \(G_{3,0,0}\), \(G_{3,1,0}\) and \(G_{3,1,1}\) respectively, where:

\[
G_{3,i,j}\left(\alpha_1, \alpha_2, \text{SNR}\right) = \frac{C_w\left(\alpha_1, \alpha_2, \lambda_{1,\text{opt}}(\alpha_1, \alpha_2, \text{SNR}), \lambda_{2,\text{opt}}(\alpha_1, \alpha_2, \text{SNR})\right)}{C_w\left(\alpha_1, \alpha_2, i, j, \text{SNR}\right)} - 1
\]

\((i,j) \in \{(0,0),(1,0),(1,1)\}\) (8).

Again, the significant hierarchical gain is \(G_3 = \min_{i,j} \left(G_{3,i,j}\right)\).

Fig. 11 represents the hierarchical gains \(G_{3,i,j}\) in terms of the SNR for \(\alpha_2 \in \{0.1, 0.25, 0.4\}\). Clearly, \(G_3\) is always very low, and it has two local maxima: the first one belonging to the first region of Fig. 10 and the second one to the third region. In addition, the first local maximum is higher than the second one. Given these results, the pertinence of a three-level hierarchy H-PAM scheme can be questioned. An exhaustive simulation of \(G_3\), performed for 0≤\(\alpha_2\)≤\(\alpha_1\)≤1 and 0≤SNR≤22 dB, led to the same conclusion.

IV. CONCLUSION

This paper highlighted the incompatibility between MRM and Shannon's definition of channel capacity, due to its exclusion of the notion of information perceptual importance, which is inherent to MRM. In order to resolve this incompatibility, a new quantity, named the “weighted channel capacity”, was defined and proposed. The simulated results of the “weighted channel capacity” in the context of H-PAM transmission on the AWGN channel indicate that this quantity could be used as an optimization criterion for choosing H-PAM hierarchical parameters.

REFERENCES


