ANALYSIS OF ADSL2's 4D-TCM PERFORMANCE

Mohamed Ghanassi, Jean François Marceau, François D. Beaulieu, and Benoît Champagne
Department of Electrical & Computer Engineering,
McGill University, Montreal, Quebec Canada
e-mail: mghanassi@tsp.ece.mcgill.ca, jfmarceau@tsp.ece.mcgill.ca, fduplessis@tsp.ece.mcgill.ca,
champagne@ece.mcgill.ca

Abstract

Asymmetric Digital Subscriber Line (ADSL) has been gaining popularity as a high speed transmission technology through the copper twisted pair telephone lines. High performance is achieved by using discrete multi-tone (DMT) modulation. DMT divides the channel into a number of independent sub-channels so that more bits are transmitted over sub-channels with higher signal-to-noise ratios. Performance can be further improved by combining DMT with trellis-coded modulation (TCM). In this paper we analyze the performance of a four-dimensional-TCM encoder as it is used in ADSL/ADSL2 modems. First, assuming all the sub-channels are transmitting the same number of bits b, we theoretically evaluate the TCM coding gain for different values of b. Then, we consider the case where sub-channels may transmit different number of bits as in an ADSL transmission. Simulation results are presented to validate our analysis.

Keywords — *Trellis-coded modulation; TCM; ADSL; ADSL2; DMT; Coding gain.*

1 Introduction

Asymmetric Digital Subscriber Line (ADSL) [1][2] is a modem technology for high-speed digital communications over copper twisted pair telephone lines. Depending on the line length and the modem capabilities, asymmetrical data rates of more than 8 Mbps downstream (to the customer) and up to 640 kbps upstream can be achieved. ADSL uses discrete multi-tone (DMT) as its modulation scheme. DMT divides the channel into a number of independent sub-channels, referred to as tones. Each tone is QAM-modulated using a different carrier. The number of bits to be transmitted in each tone is determined by a bit loading algorithm and depends on the SNR (Signal-to-Noise Ratio) of the given tone. High SNR tones carry more bits than low SNR tones.

To improve the data rate and reach performance, many features have been added in the G.992.3 standard for ADSL2 [3]. The most important addition concerns the mandatory use of a four-dimensional trellis-coded modulation (4D-TCM). This feature was previously optional in the ADSL standard.

TCM is a combined coding and modulation technique for digital transmissions over band-limited channels. It uses signal-set expansion and signal-mapping techniques to maximize the minimum Euclidian distance between coded signals. It achieves significant coding gains over uncoded modulation without compromising bandwidth efficiency. 4D-TCM coding gain consists of two components, fundamental coding gain and shaping coding gain. Fundamental coding gain is independent of the number of points in a constellation and is close to the coding gain for a high number of bits. However, some points in a finite constellation are not surrounded on all sides by other points, which affects the coding gain by a certain amount known as shaping gain. 4D-TCM performance has been reported in many papers [4][5][6][7]. In [4] and [5] an asymptotic fundamental coding gain was evaluated, whereas in [6] and [7] an error transfer function was used to theoretically calculate the fundamental coding gain.

In this paper, we evaluate the performance of a 4D-TCM scheme in a DMT modem by considering both fundamental and shaping coding gains. First, in Section 2, we theoretically evaluate lower and upper bounds of the coding gain (including fundamental and shaping gains) for different constellation sizes by considering all types of error events. Then, in Section 3, we theoretically evaluate a tight upper bound of the coding gain in an ADSL-like environment. Finally, we conclude in Section 4.

2 Trellis-coded modulation (TCM)

TCM combines redundant nonbinary modulation with a finite-state encoder which governs the selection of modulation signals, to generate coded signal sequences. Using a Viterbi algorithm, the decoder decides which of many possible sequences was most likely to have been transmitted. TCM uses signal-set expansion to provide redundancy for coding and signal-mapping functions to maximize the minimum Euclidean distance (free distance) between coded signal sequences. Signal-mapping is based on a technique called mapping by set-partitioning. It divides a signal set into smaller subsets (called "cosets") with maximally increasing the smallest distance between the subset signals. Soft decision decoding is accomplished in two steps. In the first step, called "subset decoding", for each subset of signals (characterized by parallel transitions in the code trellis), the signal closest to the received channel output is determined. In the second step, the Viterbi algorithm finds the signal path through the code trellis with the minimum sum of squared distances from the received noisy sequences. Only the signal paths already chosen by subset decoding are considered.

For QAM modulation, the constellation expansion leads to a 2-dimensional code. ADSL2 uses a 4-dimensional trellis code by concatenating two 2-dimensional QAM symbols. As shown in Fig. 1, given a pair of tones in which x and y coded bits can be transmitted, $x+y-1$ information bits $(u_{x+y-1},...,u_2,u_1)$ are extracted and coded into $x+y$ bits $(v_{x-1},...,v_1,v_0)$ and $(w_{y-1},...,w_1,w_0)$. The

TABLE I
Correspondence between 4-dimensional and 2-dimensional cosets.

4D coset	u_3	u_2	u_1	u_o	v_1	v_o	w_1	w_o	2D cosets
C_4^0	0	0	0	0	0	0	0	0	C_2^0xC_2^0
	1	0	0	0	1	1	1	1	C_2^3xC_2^3
C_4^4	0	1	0	0	0	0	1	1	C_2^0xC_2^3
	1	1	0	0	1	1	0	0	C_2^3xC_2^0
C_4^2	0	0	1	0	1	0	1	0	C_2^2xC_2^2
	1	0	1	0	0	1	0	1	C_2^1xC_2^1
C_4^6	0	1	1	0	1	0	0	1	C_2^2xC_2^1
	1	1	1	0	0	1	1	0	C_2^1xC_2^2
C_4^1	0	0	0	1	0	0	1	0	C_2^0xC_2^2
	1	0	0	1	1	1	0	1	C_2^3xC_2^1
C_4^5	0	1	0	1	0	0	0	1	C_2^0xC_2^1
	1	1	0	1	1	1	1	0	C_2^3xC_2^2
C_4^3	0	0	1	1	1	0	0	0	C_2^2xC_2^0
	1	0	1	1	0	1	1	1	C_2^1xC_2^3
C_4^7	0	1	1	1	1	0	1	1	C_2^2xC_2^3
	1	1	1	1	0	1	0	0	C_2^1xC_2^0

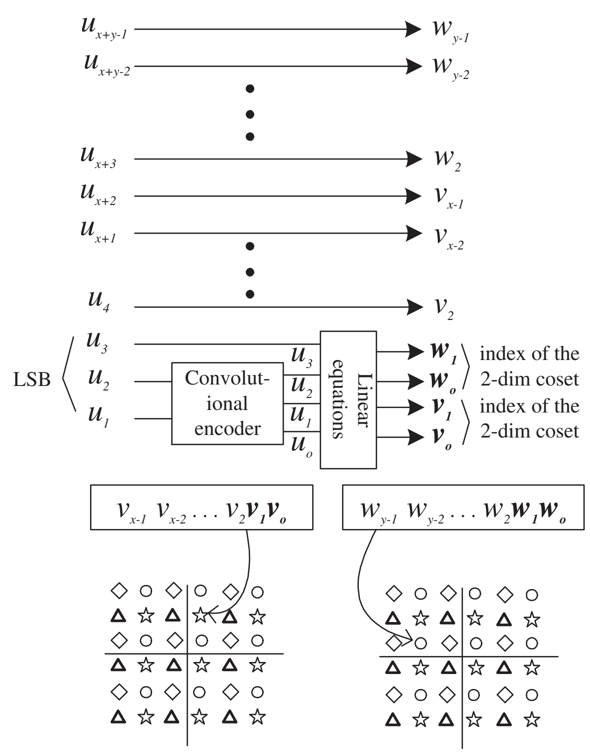

Figure 1. Trellis coding in ADSL/ADSL2.

three least significant bits (LSB) (u_3, u_2, u_1) determine the bits (v_1, v_0) and (w_1, w_0) which are the 2 least significant bits of a constellation point. These bits (shown in bold in Fig. 2) are the binary representation of the index of the 2-dimensional cosets in which the constellation point lies. u_o is the result of encoding (u_2, u_1), while (v_1, v_0) and (w_1, w_0) are computed from (u_3, u_2, u_1, u_0) using linear equations.

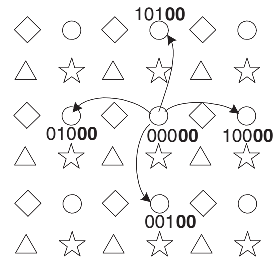

Figure 2. Error event of type 1 (parallel transition in the same subset).

Table I shows the relation between 4-dimensional and 2-dimensional cosets.

2.1 Performance evaluation - Coding gain

In our performance evaluation, we consider an uncoded system with a minimum Euclidean distance d_u. Since TCM adds 1 bit per pair of tones, half of the tones have their constellation doubled. In order to keep the same transmitted signal power, constellations of the coded signal have to be scaled down before transmission. As a result, the minimum Euclidian distance is reduced and can be written as $d_c = \gamma d_u$, where $\gamma < 1$.

In order to determine the coding gain, different error events have to be considered. In the 4D-TCM there are three types of errors. The first type of error occurs when the received symbol **v** or **w** is closer to a symbol from the same subset but different from the symbol transmitted. This parallel transition is illustrated in Fig. 2; the transmitted symbol is 000**00** but the received symbol is 101**00**, 010**00**, 001**00**, or 100**00**. The squared Euclidean distance for this error type is $(2d_c)^2 = 4d_c^2$.

The second type of error is illustrated in Fig. 3. It occurs when the received symbols **v** and **w** are both closer to symbols of the same 4-D coset as for the transmitted symbols but not of the same 2-D subsets. For example, in Fig. 3, the transmitted and received symbols are both from the C_4^4 4D-coset. For **v**, the transmitted symbol is 000**11** (from the C_2^3 subset), whereas the received symbol is 010**00**, 001**00**, 000**00**, or 011**00** (from the C_2^0 subset). For **w**, the transmitted symbol is 000**00** (from the C_2^0 subset) whereas the received symbol is 000**11**, 001**11**, 011**11**, or 010**11** (from the C_2^3 subset). According to Table I, this error event corresponds to an error on bit u_3. The squared Euclidean distance for this error is $2d_c^2 + 2d_c^2 = 4d_c^2$.

The third type of error event is related to a trellis sequence. It occurs when a path in the trellis diagram diverges from the true path and remerges after a few stages as illustrated in Fig. 4. By analyzing the TCM state diagram, we can show that the minimum squared distance between branches that diverge from or converge to a same state is $2d_c^2$. Since two paths diverge and remerge after at least three stages, the minimum squared Euclidean distance between true and erroneous paths is $2d_c^2 + d_c^2 + 2d_c^2 = 5d_c^2$.

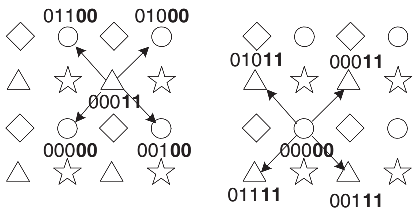

Figure 3. Error event of type 2 (symbols of different subsets within the same 4-D coset).

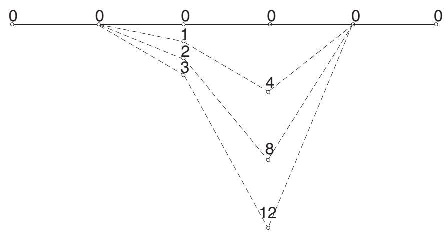

Figure 4. Paths in the trellis that diverge in one state and remerge in another.

For an additive white Gaussian noise (AWGN) channel, we can establish a union bound P_{up} on the probability of a symbol error as

$$P_{up} = N_{s1}N_{p1}Q\Big(\frac{\gamma d_u}{\sigma}\Big) + \frac{1}{2}N_{s2}N_{p2}Q\Big(\frac{\gamma d_u}{\sigma}\Big) + \\ + \frac{1}{2}\sum_{\beta=0}^{\infty} N_{s3}(\beta)N_{p3}(\beta)Q\Big(\frac{\gamma\beta d_u}{2\sigma}\Big) \qquad (1)$$

where N_{pi} is the average number of paths for errors of type $i(i=1,2)$, $N_{p3}(\beta)$ is the average number of paths for error of type 3, which depends on a factor β, N_{si} is the average number of symbols in error associated to an error event of type $i(i=1,2)$, and $N_{s3}(\beta)$ is the average number of symbols in error associated to an error event of type 3. β is the Euclidean distance for error of type 3, normalized by d_c. σ^2 is the noise variance. $Q(x)$ is related to the complementary error function $erfc(x)$ by $Q(x) = 0.5erfc(x/\sqrt{2})$. We calculated N_{pi} and N_{si} using the transfer function of the convolutional code for different number of bits. Results are shown in Tables II and III. We obtained the same results with another method by "walking" into the trellis and retaining all the possible error event paths.

The TCM coding gain can be bounded by lower and upper values. The lower bound is obtained by considering all error events for the coded system and by using the probability of error P_{up} defined in (1). The upper bound for coding gain is derived by considering only error events of type 1 and 2 with the least Euclidian distance (which are dominant at high SNR). Hence, the corresponding lower bound of the probability of a symbol error is

$$P_{low} = N_{s1}N_{p1}Q\Big(\frac{\gamma d_u}{\sigma}\Big) + \frac{1}{2}N_{s2}N_{p2}Q\Big(\frac{\gamma d_u}{\sigma}\Big) \qquad (2)$$

TABLE II
AVERAGE NUMBER OF PATHS (N_{pi}) FOR DIFFERENT TYPES OF ERROR EVENTS.

Error type	1	2	3				
β^2	4	4	5	6	7	8	9
Number of bits b	Number of paths						
2	0.00	1.00	16	88	416	2008	10128
3	1.00	2.25	23	122	602	3037	15782
4	2.00	5.06	118	958	6816	49541	375880
5	2.50	8.27	193	1726	13347	105562	870033
6	3.00	9.38	254	2452	20235	170272	1499039
7	3.25	11.39	319	3182	27319	239629	2194465
8	3.50	12.36	370	3823	33892	306742	2900915
9	3.63	13.37	404	4241	38237	352043	3386145
10	3.75	14.09	436	4655	42636	398746	3896718
11	3.81	14.59	455	4890	45155	425763	4194672
12	3.88	15.02	473	5124	47690	453198	4500273
13	3.91	15.27	482	5248	49044	467937	4665178
14	3.94	15.51	492	5373	50403	482784	4832116
15	3.95	15.63	497	5437	51105	490477	4918799
∞	4	16	512	5632	53248	514048	5191680

TABLE III
AVERAGE NUMBER OF SYMBOLS IN ERROR (N_{si}) FOR DIFFERENT TYPES OF ERROR EVENTS.

Error type	1	2	3				
β^2	4	4	5	6	7	8	9
N_{si}	1.00	2.00	4.25	5.27	6.20	7.11	8.04

The TCM encoder uses three information bits from each pair of tones to provide four coded bits. If we consider an uncoded system with the same number of bits for all the tones, the coded system will have one more bit for half of the tones. Since a probability of error for tones with a different number of bits is difficult to derive, it is more convenient to consider a coded system with the same number of bits for each tone. Therefore, for the corresponding uncoded system, half of these tones have one less bit. In this case, the probability of error for a coded system is given by (1) and the probability of error for an uncoded system is well approximated by

$$P_{unc} = \frac{1}{2}\Big[N_{unc}(b) + N_{unc}(b-1)\Big]Q\Big(\frac{d_u}{2\sigma}\Big) \qquad (3)$$

where $N_{unc}(b)$ is the average number of neighbors at the minimum Euclidean distance, shown in Table IV. Table V shows the average signal energy for coded and uncoded constellations of different sizes, as well as the related quantity $\gamma^2 = d_c^2/d_u^2$.

The probability of a symbol error for an uncoded system and the upper and lower bounds for a coded system as given by (1) and (2) are plotted in Fig. 5 for $b = 6$. Simulation

TABLE IV

AVERAGE NUMBER (N_{unc}) OF NEIGHBORS FOR QAM CONSTELLATIONS.

Nb. of bits	2	3	4	5	6	7	8
N_{unc}	2.00	2.00	3.00	3.25	3.50	3.63	3.75
Nb. of bits	9	10	11	12	13	14	15
N_{unc}	3.81	3.88	3.91	3.94	3.95	3.97	3.98

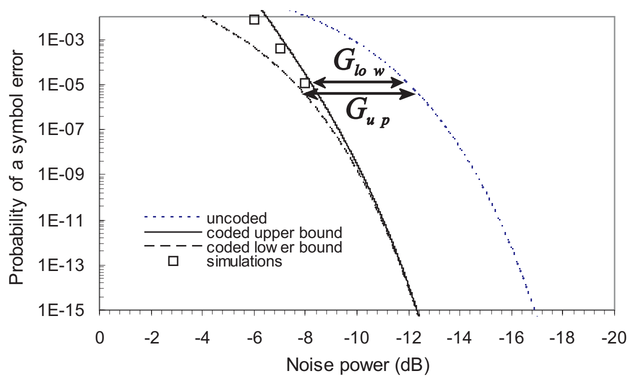

Figure 5. Probabilities of error for uncoded and coded systems. $G_{up}(G_{low})$ is the upper(lower) bound of the coding gain.

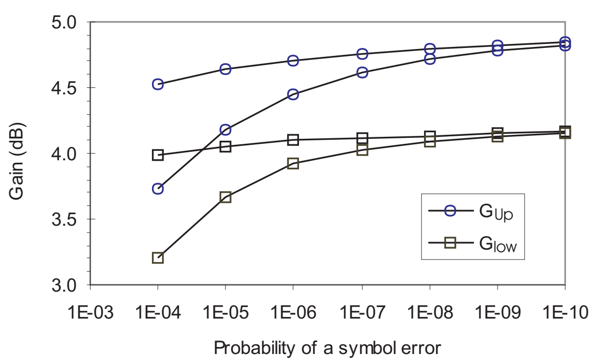

Figure 6. Minimum and maximum values of G_{up} (circles) and G_{low} (squares).

results, also shown, are close to the upper bound. For a specific probability of error, the coding gain is calculated as the noise power difference between coded and uncoded systems (see Fig. 5).

Upper and lower bounds of the coding gain, G_{up} and G_{low}, are calculated for different number of bits and for coded probabilities of error between 10^{-4} and 10^{-10}. For each probability of error, we determine the minimum and maximum values of G_{low} and G_{up} by considering all possible values of b (except $b = 2$ since in this case the factor γ, equal to zero, is quite different from the values corresponding to $b = 3$ to 15). As shown in Fig. 6, the lower bound

TABLE V

ENERGY INCREASE FOR CODED SIGNAL. $E_{cod}(E_{unc})$ IS THE AVERAGE ENERGY OF CODED (UNCODED) SIGNAL.

Nb. of bits	E_{cod}	E_{unc}	$E_{cod}/E_{unc} = 1/\gamma^2$	$\gamma^2(dB)$
2	4	4	1.00	0.00
3	12	8	1.50	−1.76
4	20	16	1.25	−0.97
5	40	30	1.33	−1.25
6	84	62	1.35	−1.32
7	164	124	1.32	−1.21
8	340	252	1.35	−1.30
9	660	500	1.32	−1.21
10	1364	1012	1.35	−1.30
11	2644	2004	1.32	−1.20
12	5460	4052	1.35	−1.30
13	10580	8020	1.32	−1.20
14	21844	16212	1.35	−1.29
15	42324	32084	1.32	−1.20

of the coding gain at a probability of error of 10^{-4} is at least 3.2 dB. For low probabilities of error (or high SNR), the minimum and maximum of G_{low} and G_{up} converge to 4.2 dB for G_{low} and 4.8 dB for G_{up}. In other words, the coding gain is bounded between 4.2 and 4.8 dB, which is in very good agreement with the theoretical asymptotic coding gain value of 4.5 dB calculated from the free distance gain [4].

3 DMT-ADSL

In a DMT modulation, the transmission channel is partitioned into parallel independent narrowband subchannels. Each subchannel transmits a quadrature amplitude modulated signal. In a DSL environment, the signal-to-noise ratio (SNR) is frequency-dependent and the number of bits b allocated to each subchannel depends on its SNR as follows

$$b = log_2\Big(1 + \frac{SNR \cdot CG}{\Gamma \cdot M}\Big) \tag{4}$$

where CG is a coding gain associated to a coding scheme, Γ is the Shannon limit of 9.75 dB corresponding to a probability of error of 10^{-7} for QAM modulation, and M is a noise margin. For ADSL2, bit loading is also performed according to (4). However, in order to transmit the same information bit rate as ADSL, ADSL2 loads one more bit for each pair of tones. This is accomplished by reducing the noise margin by approximately 1.5 dB. Using typical values of 3 dB for CG and at least 6 dB for M in the bit loading process of (4), the probability of a symbol error can be much lower than 10^{-7}. Hence, in these conditions we can consider that error events of types 1 and 2 are dominant and the coding gain for an ADSL2 transmission can be estimated from Fig. 6, i. e. between 4.2 dB and 4.8 dB.

In a DMT-ADSL configuration, the probability of a symbol error for an uncoded system can be written as

$$P_{unc} = \frac{1}{N}\sum_{i=0}^{N} N_{unc}(b_i) Q\Big(\frac{d_{u,i}}{\sigma_i}\Big) \tag{5}$$

where N is the number of tones, b_i is the number of bits for tone i, σ_i^2 is the noise variance for tone i, $d_{u,i}$ is the minimum Euclidian distance for the uncoded constellation corresponding to tone i, and N_{unc} is the average number of neighbors at the minimum Euclidean distance as shown in Table IV.

For a coded system, calculating the theoretical probability of a symbol error by considering all types of errors is too complex. However, by considering only error events of types 1 and 2, we can establish a theoretical lower bound for the probability of a symbol error:

$$P_{low} = \frac{2}{N}\sum_{i=0}^{N/2-1} \sqrt{N_{(2i)1}N_{(2i+1)1}}\, Q\Big(\sqrt{\frac{d_{c,i}^2 + d_{c,i+1}^2}{\sigma_i^2+\sigma_{i+1}^2}}\Big) + \\ + \frac{1}{N}\sum_{i=0}^{N-1} N_{i2} Q\Big(\frac{d_{c,i}}{\sigma_i}\Big) \tag{6}$$

where N_{ik} is the number of paths for error events of type $k (k = 1, 2)$ for tone i, $d_{c,i}$ is the minimum Euclidian distance of the coded constellation corresponding to tone i.

In order to verify our theoretical analysis, we simulated an ADSL transmission for a loop of 9 kilofeet, with a data rate of 4 Mbps, in an environment comprising 49 pairs of HDSL disturbers. In order to vary the SNR for each tone by the same amount, we attenuated the transmitted DMT signal. The symbol error rate (SER) was then measured.

First, to verify the amount of power penalty due to the constellation expansion introduced by coding, we simulated an uncoded system with a tone configuration corresponding to the specified transmission conditions, and we also simulated an uncoded system with the same configuration but with 1 more bit for each pair of tones. Simulation results perfectly matched the theoretical results (calculated by (5)). As shown in Fig. 7, the x-axis difference between the two uncoded curves is 1.5 dB, which corresponds to the power penalty due to constellation expansion.

Simulation results for coded system, shown in Fig. 7, are close to the theoretical lower bound. For a probability of error of 10^{-4}, the experimental coding gain is around 3.5 dB, while for lower probability of errors, it converges to a 4.5 dB limit.

4 Conclusion

In this paper we have analyzed the performance of a four dimensional-TCM encoder as it is used in ADSL/ADSL2 modems. We considered all types of error events and theoretically evaluated lower and upper bounds of the coding gain for different constellation sizes. Simulation results in

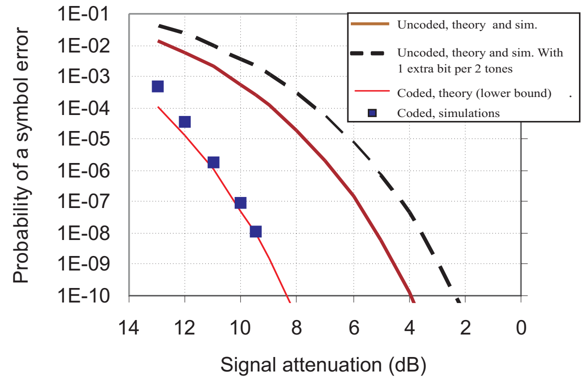

Figure 7. SER simulation results for an ADSL transmission.

an AWGN environment showed good agreement with theory. TCM performance was also evaluated for an ADSL2 modem in a DSL environment. Simulation results for the probability of a symbol error, as well as for the coding gain are in good agreement with theory.

Acknowledgments

Support for this work was provided by a joint research grant from the Natural Sciences and Engineering Research Council of Canada (NSERC) and Bell Canada.

References

[1] American National Standard T1.413-1998, *"Network and Customer Installation Interfaces - Asymetric Digital Suscriber Line (ADSL) Metallic Interface,"* 1998.

[2] ITU-T Recommandation G.992.1, *"Asymetric Digital Suscriber Line (ADSL) Tranceivers),"* 1999.

[3] ITU-T Recommandation G.992.3, *"Asymetric Digital Suscriber Line (ADSL) Tranceivers 2 - (ADSL2),"* March 2002.

[4] L. F. Wei, *"Trellis-Coded Modulation with Multidimensional Constellation,"* IEEE Trans. Inform. Theory. vol. 33, no. 4, pp. 483-501, Jul. 1996.

[5] G. Ungerboeck, *"Trellis-Coded Modulation with Redundant Signal Sets: Parts I and II," IEEE communications magazine.* vol. 25, pp. 12-21, Feb. 1987.

[6] T. N. Zogakis, J. T. Aslanis Jr. and J. M. Cioffi, *"Analysis of a concatenated coding gain scheme for a discrete multitone modulation system," IEEE Military Communication Conf.*, pp. 433-437, Oct. 1994.

[7] M. Cao, K. R. Subramanian and V. K. Dubey, *"Multidimensionnal TCM schemes for ADSL," Electronic Letters*, vol. 35, no. 11, pp. 870-872, June 1999.