Hierarchical Generalized Cantor Set Modulation

Simon Görtzen, Lars Schiefler, Anke Schmeink

Information Theory and Systematic Design of Communication Systems
UMIC Research Centre, RWTH Aachen University, 52056 Aachen, Germany
Email: {goertzen, schiefler, schmeink}@umic.rwth-aachen.de

Abstract—In this paper, we show that arbitrary hierarchical pulse amplitude modulation (PAM) schemes can be fully described by generalized Cantor sets. Generalized Cantor sets are modified versions of the Cantor ternary set, a famous mathematical construct known for its set-theoretical properties. The fractal nature of generalized Cantor sets allow for a natural reinterpretation as a modulation scheme. The resulting Cantor set description of one-dimensional hierarchical modulation schemes covers the constellation points as well as the boundary points of the decision regions. Furthermore, we derive simple formulas for the average signal power as well as for iterative demodulation. All results can be extended to two dimensions and hierarchical quadrature amplitude modulation (QAM) schemes. As such, this paper offers a novel perspective on the classification and parametrization of practical hierarchical modulation schemes.

I. Introduction

In his information-theoretic work on broadcast channels, Cover [1] shows that it is possible to outperform time-sharing strategies by superimposing high-rate information on low-rate information. The high-rate, low-priority information is recovered by receivers with a high signal-to-noise ratio (SNR), but appears as noise to receivers in low SNR environments. Despite this, those receivers still recover the low-rate, high-priority information. These insights have motivated further research in the area of unequal error protection [2], [3] with applications, for example, in the field of digital video broadcasting [4]–[6]. To achieve unequal error protection in practice, hierarchical signal constellations are employed, specifically hierarchical quadrature amplitude modulation (QAM) [7]–[10]. By applying nonuniform signal spacing, the most significant bits experience much better error protection than others. To evaluate the performance of hierarchical QAM, most research focuses on the computation of the bit error rate (BER) of certain hierarchical signal constellations. These computations are usually performed for systems with a small number of information layers and parametrized by distances obtained from the constellation diagram.

In this paper, we present a constructive approach to obtain arbitrary hierarchical QAM constellations. We show that the constellation points of hierarchical QAM coincide with the elements of sets that arise naturally in the construction of generalized Cantor sets. These are modified versions of the famous Cantor ternary set, a mathematical construct known mainly for its set-theoretical properties. We show that by exploiting their fractal nature, Cantor sets can be used to

describe generalized Gray-coded hierarchical pulse amplitude modulation (PAM) schemes. This includes a Cantor-based description of the decision regions of hierarchical PAM. The extension to two dimensions then yields corresponding results for square and rectangular hierarchical QAM constellations.

The remainder of this paper is organized as follows. In Section II, we define the real-valued channel model and present the basics of PAM and hierarchical PAM. In Section III, we present generalized Cantor sets and their properties. Section IV introduces the Cantor set modulation scheme. The extension to two dimensions is presented in Section V. Following this section, we present exemplary symbol error rate computations and analyze how path loss effects naturally enable hierarchical information transmission in Section VI. Section VII concludes the paper.

II. The AWGN Channel and PAM

Consider the reception of a broadcasted real-valued symbol X through additive white Gaussian noise (AWGN) with power spectral density $N_0/2$. If we assume perfect channel state information at the receiver, the received signal Y is

$$Y = X + W, \tag{1}$$

with $W \sim \mathcal{N}(0, N_0/2)$. The average symbol energy is defined as $E_s = \mathbb{E}(X^2)$. Assuming X to be uniformly distributed between symbols x_i, $i = 1, \ldots, M$, we obtain

$$E_s = \frac{1}{M} \sum_{i=1}^{M} x_i^2. \tag{2}$$

With this, the average SNR γ at the receiver can be expressed in terms of the signal energy per symbol E_s divided by N_0,

$$\gamma = \frac{E_s}{N_0}. \tag{3}$$

The constellation for PAM is designed to minimize the symbol error rate when transmitting over the real-valued AWGN channel. This is achieved by an equidistant distribution of the constellation points, centered at the origin. The average symbol error rate P_s is a function of γ and can be expressed as

$$P_s(\gamma) = \frac{2(M-1)}{M} Q\left(\sqrt{\frac{6\gamma}{M^2 - 1}}\right), \tag{4}$$

where $Q(x) = 1 - \Phi(x)$ describes the tail probability of the standard normal distribution. A detailed description can be
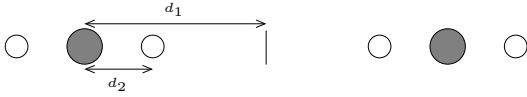
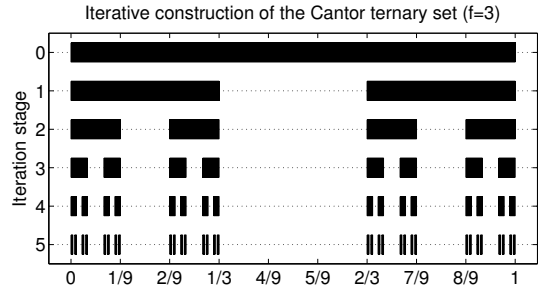Fig. 1. Hierarchical 2/4-PAM. The larger circles denote the virtual symbols.

Fig. 2. This figure shows the iterative construction of the Cantor set for the special case $f_i = 3$, $i \in \mathbb{N}$. Shown are the Cantor sets of stage 0 to 5.

found in [11]. We denote an equidistant PAM scheme with M symbols by M-PAM.

Next, we give a short description of hierarchical PAM. To achieve unequal error protection, the decision regions for low-rate information are enlarged, while the decision regions for high-rate information are reduced in size. For users with low SNR, the symbols carrying high-rate information are indistinguishable from each other and the modulation appears to be an N-PAM scheme with $N < M$. Note however, that these N virtual symbols are not part of the constellation diagram and only appear at the receiver side when interpreting the high-rate information as noise. We denote hierarchical PAM with k layers of information and N_i, $i = 1, \ldots, k$, (virtual) symbols in each layer by $N_1 / \ldots / N_k$-PAM. See Figure 1 for an exemplary 2/4-PAM scheme.

To describe a general $N_1 / \ldots / N_k$-PAM scheme, parameters $d_1, \ldots, d_k$ are required. They are computed from the constellation diagram as follows: For $i = 1, \ldots, k$, a receiver recovering i-th layer information experiences an $N_1 / \ldots / N_i$-PAM system. Except for $i = k$, the symbols N_i are virtual ones. The minimum distance between two symbols in N_i is defined to be $2d_i$.

Refer to Figure 1 for the significance of d_1 and d_2 in the 2/4-PAM case. Clearly, $2d_2 \le d_1$ has to hold to achieve reasonable error protection. For $2d_2 = d_1$, the system is equivalent to 4-PAM, and for $d_2 = 0$ it degrades to 2-PAM.

III. THE GENERALIZED CANTOR SET

The Cantor set is a mathematical construct named after Georg Cantor and famous for its set-theoretical and topological properties. We give a short definition taken from [12]. The Cantor set consists of all points in the closed unit interval which can be expressed to the base 3 without using the digit 1. This ternary representation is the reason why it is also called Cantor ternary set. Geometrically, the Cantor set is obtained by deleting a sequence of middle thirds from the closed unit interval $[0, 1]$. First, $(1/3, 2/3)$ is removed, leaving $[0, 1/3] \cup [2/3, 1]$. In the next step, the intervals $(1/9, 2/9)$ and $(7/9, 8/9)$ are deleted. This process of removing the middle third of each interval is continued ad infinitum and the remaining points make up the Cantor set. See Figure 2 for a visualization of this iterative process.

Motivated by the Cantor ternary set, we now define generalized Cantor sets. In a preliminary step, we center the set around the origin. Therefore, we start with the interval $[-1, 1]$ instead of the unit interval. The next step is to allow not only the middle third of each interval to be removed, but arbitrary ratios. We introduce a scaling factor f such that from each interval I, the middle portion of size $|I|(f-2)/f$ is deleted.

Clearly, the choice $f = 3$ yields the Cantor ternary set. Finally, we allow the use of different factors f_i at every stage of the iteration.

We now formally define the generalized Cantor set C on the interval $[-1, 1]$ subject to scaling factors $(f_i)_{i \in \mathbb{N}}$ with $f_i \ge 2$, $i \in \mathbb{N}$. Let C_n denote the Cantor set of stage n, $n \in \mathbb{N}_0$. Starting with $C_0 = [-1, 1]$, the Cantor set of stage i is obtained by removing a portion from the middle of each interval I in C_{i-1}, such that two intervals of size $|I|/f_i$ remain. For the special case $f_i = 2$, it holds that $C_i = C_{i-1}$ as an empty interval is removed, but we still regard this step as a further division into two equidistant intervals. We obtain the iterative definition

$$C_0 = [-1, 1], \tag{5}$$

$$C_i = \frac{1}{f_i}(C_{i-1} - (f_i - 1)) \cup \frac{1}{f_i}(C_{i-1} + (f_i - 1)) \tag{6}$$

for all $i \ge 1$, in which we use the notation

$$a \cdot S + b = \{as + b \mid s \in S\} \tag{7}$$

for a set S and $a, b \in \mathbb{R}$. Clearly, C_n is a union of 2^n intervals of length $2/\prod_{i=1}^{n} f_i$. The generalized Cantor set itself is defined as

$$C = \lim_{n \to \infty} C_n, \tag{8}$$

however, for obvious practical reasons, we focus on Cantor sets of finite stage in this paper. Note that the generalized Cantor set can be extended to multiple dimensions. In particular, $C \times C$ denotes the two-dimensional Cantor set, which is generally referred to as Cantor dust. Consequently, $C_n \times C_n$ describes the Cantor dust of stage n.

Following this definition, we analyze particular discrete subsets of stage-n Cantor sets. Let P_n denote the set of centers of all intervals of C_n. Clearly, $|P_n| = 2^n$ holds. Furthermore, the family $(P_n)_{n \in \mathbb{N}_0}$ can be represented through a binary tree, in which the nodes at depth n correspond to P_n. With the description

$$P_0 = \{0\}, \tag{9}$$

$$P_i = \left(P_{i-1} - \frac{f_i - 1}{\prod_{j=1}^{i} f_j}\right) \cup \left(P_{i-1} + \frac{f_i - 1}{\prod_{j=1}^{i} f_j}\right) \tag{10}$$

each node x at depth $n-1$ has offspring $x \pm (f_n - 1)/\prod_{i=1}^{n} f_i$ and all levels of the tree are ordered from left to right. This binary tree representation allows to map each element of P_n to a bit sequence of length n. We derive a closed-form expression for the average power of a real-valued symbol X_n taking on the values P_n.

Proposition 1. *For $n \in \mathbb{N}_0$, let X_n be a uniformly distributed random variable with support P_n, i.e., $P(X_n = p) = \frac{1}{2^n}$ for all $p \in P_n$. Then, the average power of X_n computes to*

$$\mathbb{E}(X_n^2) = \sum_{j=1}^{n} \left(\frac{f_j - 1}{\prod_{i=1}^{j} f_i} \right)^2. \tag{11}$$

Proof: It holds that $\mathbb{E}(X_0^2) = 0$ and

$$\mathbb{E}(X_n^2) = \frac{1}{2^n} \sum_{p \in P_n} p^2 \tag{12}$$

$$= \frac{1}{2^n} \sum_{p \in P_{n-1}} \left(p - \frac{f_n - 1}{\prod_{i=1}^{n} f_i} \right)^2 + \left(p + \frac{f_n - 1}{\prod_{i=1}^{n} f_i} \right)^2 \tag{13}$$

$$= \frac{1}{2^{n-1}} \sum_{p \in P_{n-1}} p^2 + \left(\frac{f_n - 1}{\prod_{i=1}^{n} f_i} \right)^2 \tag{14}$$

$$= \mathbb{E}(X_{n-1}^2) + \left(\frac{f_n - 1}{\prod_{i=1}^{n} f_i} \right)^2 = \sum_{j=1}^{n} \left(\frac{f_j - 1}{\prod_{i=1}^{j} f_i} \right)^2. \tag{15}$$

∎

This result concludes the section on the generalized Cantor set. In the following section we show how the fractal nature of generalized Cantor sets naturally motivates hierarchical modulation schemes.

IV. CANTOR SET MODULATION

Let $n \in \mathbb{N}$ denote the number of bits to be transmitted, with corresponding scaling factors $f_1, \ldots, f_n \geq 2$. The transmission of an arbitrary bit sequence $(b_1, \ldots, b_n) \in \{0,1\}^n$ is realized through the following steps. To reduce the overall bit error rate, the sequences corresponding to two adjacent symbols are only allowed to differ by one bit. This is easily accomplished with Gray coding [13]. The Gray coded sequence is computed with an exclusive-or operation ($\oplus$) and given by $(b_1', \ldots, b_n')$ as

$$b_1' = b_1, \tag{16}$$

$$b_{i+1}' = b_i' \oplus b_{i+1}. \tag{17}$$

Through the binary tree representation of P_n, we obtain a one-to-one mapping

$$\{0,1\}^n \to P_n, \tag{18}$$

$$(b_1, \ldots, b_n) \mapsto \sum_{i=1}^{n} (-1)^{1-b_i'} \frac{f_i - 1}{\prod_{j=1}^{i} f_j}. \tag{19}$$

Therefore, the constellation points of the n-bit Cantor set modulation are the elements of P_n. At the receiver side,

iterative decoding of the received value $Y = y_1$ is performed as follows:

$$s_i = \operatorname{sgn}(y_i), \tag{20}$$

$$y_{i+1} = f_i y_i - s_i (f_i - 1), \tag{21}$$

$$c_i' = \frac{1}{2}(1 + s_i), \tag{22}$$

$$c_1 = c_1', \tag{23}$$

$$c_{i+1} = c_i \oplus c_{i+1}. \tag{24}$$

Thus, $(c_1, \ldots, c_n)$ denotes the received bit sequence. In the case of an error-free reception, $b_i = c_i$, $i = 1, \ldots, n$, holds. The iterative decoding in (21) is made possible by the fractal nature of the generalized Cantor set. We traverse the nodes of the binary tree containing the elements of P_n at depth n and at each step decide whether to go left ($c_i' = 0$) or right ($c_i' = 1$) based on the sign of y_i. Despite the complexity of the Cantor set, it is possible to exactly describe the decision regions of this modulation scheme. Those are the intervals bounded by the points

$$\mathcal{D} = \bigcup_{i=0}^{n-1} P_i \cup \{\pm\infty\}. \tag{25}$$

Furthermore, the elements of P_{i-1} are the boundary points of the decision regions which are relevant for the decoding of b_i, the i-th most significant bit.

We call a modulation of this type Cantor set pulse amplitude modulation, or CPAM for short. The notation $CPAM(f_1, \ldots, f_n)$ is used to describe the choice of scaling factors. Clearly, this modulation allows for almost arbitrary signal constellations while maintaining a simple iterative demodulation procedure. In fact, CPAM can be used to holistically describe arbitrary hierarchical PAM schemes as introduced in Section II.

Proposition 2. *For $n \in \mathbb{N}$, the constellation points of PAM with $M = 2^n$ coincide with the constellation points of $CPAM(f_1, \ldots, f_n)$ with $f_1 = \cdots = f_n = 2$. The same is true for the decision regions. Thus, PAM is a special case of CPAM with $PAM = CPAM(2, \ldots, 2)$.*

This result extends to hierarchical PAM schemes of arbitrary size. Let $d_1, \ldots, d_k$ describe the distance parameters of a full $N_1/\ldots/N_k$-PAM scheme with $N_i = 2^i$, $i = 1, \ldots, k$. Then,

$$N_1/\ldots/N_k\text{-}PAM = CPAM(f_1, \ldots, f_k) \tag{26}$$

with

$$f_k = 2, \tag{27}$$

$$f_i = \frac{d_i}{d_{i+1}} \frac{f_{i+1} - 1}{f_{i+1}} + 1, \tag{28}$$

for $i = 1, \ldots, k-1$.

Proof: From (10) we conclude that the distances d_i have to coincide with

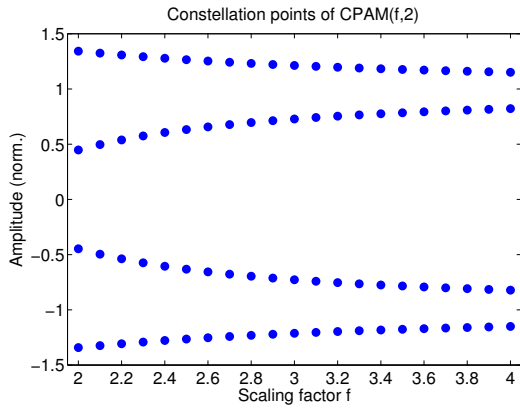$$F_i = \frac{f_i - 1}{\prod_{j=1}^{i} f_j} \tag{29}$$

Fig. 3. Sample of constellation points of CPAM$(f, 2)$. Note the equidistant 4-PAM spacing at $f = 2$ and the slow convergence towards BPSK (2-PAM).
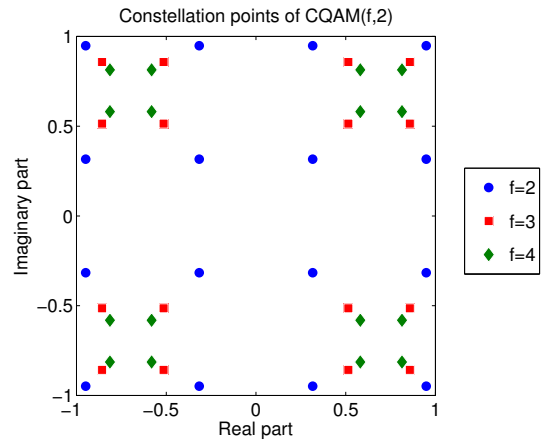

Fig. 4. Sample of constellation points of CQAM$(f, 2)$. Note the 16-QAM spacing for $f = 2$ and the slow convergence towards QPSK (4-QAM).

up to a constant factor. Thus, (26) is equivalent to

$$\frac{F_i}{F_{i+1}} = \frac{d_i}{d_{i+1}} \tag{30}$$

for $i = 1, \ldots, k - 1$. This is easily verified with (28):

$$\frac{F_i}{F_{i+1}} = (f_i - 1) \frac{f_{i+1}}{f_{i+1} - 1} \tag{31}$$

$$= \frac{d_i}{d_{i+1}} \frac{f_{i+1} - 1}{f_{i+1}} \frac{f_{i+1}}{f_{i+1} - 1} = \frac{d_i}{d_{i+1}}. \tag{32}$$

∎

For an example of this equivalence, refer to Figure 3, which depicts the constellation points of CPAM$(f, 2)$ for varying values of the scaling factor f. Here, a 2-bit transmission is modulated in which the sign of the symbol determines the value b_1 of the first bit of information. For $f = 2$, the CPAM constellation coincides with regular PAM. For increasing f, the distance between points carrying different values of b_1 increases, while the distance between points with the same first bit decreases. Thus, the reception of the first bit is simplified at the cost of a more complicated reception of the whole symbol.

V. CANTOR DUST QAM

To extend PAM with constellation size M to QAM, the real-valued symbols are placed into the complex plane in a square fashion. The resulting modulation scheme has M^2 distinct symbols with an average energy per symbol which is twice as high as for PAM. This is equivalent to a halving of the SNR γ computed for PAM. The extension of CPAM to Cantor dust QAM (CQAM) follows the same rules and results in a similar formula for the average symbol error probability. Let $P_s^{\text{CPAM}}(\gamma)$ denote the average symbol error rate of CPAM at SNR γ. Then, the average symbol error rate in two dimensions is computed as

$$P_s^{\text{CQAM}}(\gamma) = 1 - \left(1 - P_s^{\text{CPAM}}\left(\frac{\gamma}{2}\right)\right)^2. \tag{33}$$

See Figure 4 for a sample of the CQAM$(f, 2)$ constellation points for different values of f. Here, the quadrant of the
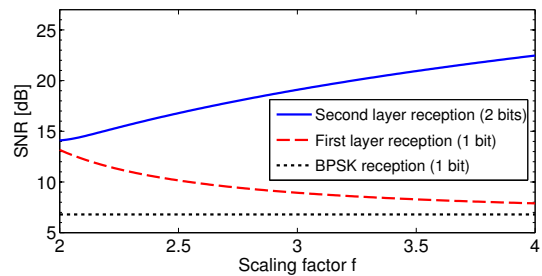

Fig. 5. Required SNR for CPAM$(f, 2)$ transmission at $P_s^{\text{CPAM}} = 10^{-3}$.

symbol carries the information of the first two bits. For increasing f, this information is protected at the cost of a more complicated reception of the whole symbol.

For non-square constellations, the extension to two-dimensional rectangular CQAM is performed analogously. We omit the details.

VI. HIERARCHICAL INFORMATION TRANSMISSION

In this section, we investigate the symbol error rates of a 2-bit transmission with two layers of information, i.e., 2/4-PAM. Correct reception is assumed as long as the average symbol error rate does not exceed a target P_s. We show that by applying a CPAM$(f, 2)$-modulation, the required SNR for the first layer can be decreased while the required SNR for the second layer increases compared to a regular 4-PAM scheme. See Figure 5 for the numerical results for varying scaling factors f.

Next, we cover the extension of this example to two dimensions, i.e., 4/16-QAM. As the number of symbols is squared, the number of transmitted bits in each layer doubles. This means that the first layer of CQAM transmission includes two bits of information, while the second layer transmission includes four. From (33), it is possible to compute the required SNR for a successful CQAM$(f, 2)$-transmission. See Figure 6 for the results. Evidently, the extension to two dimensions
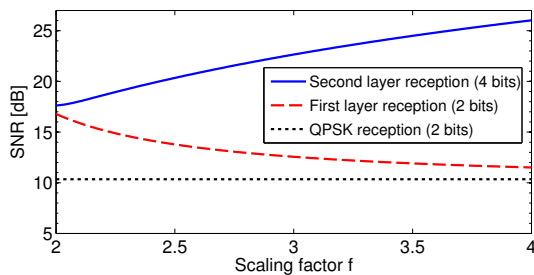
Fig. 6. Required SNR for CQAM$(f, 2)$ transmission at $P_s^{\mathrm{CQAM}} = 10^{-3}$.

Fig. 7. Trade-off between distance quotient $\rho = D_2/D_1$ and additional energy required for CQAM$(f, 2)$ at $P_s^{\mathrm{CQAM}} = 10^{-3}$.

has no significant influence on the general behavior of the modulation scheme.

In a final step we analyze how path loss naturally enables the broadcasting of hierarchical information. We assume a simple path loss model with a path loss exponent α such that the SNR γ at the receiver is proportional to $D^{-\alpha}$, with D denoting the distance between the transmitter and the receiver. We denote the maximal distance for layer one and layer two reception by D_1 and D_2, respectively. Note that this implies $D_2 \leq D_1$. Thus, the quotient $\rho = D_2/D_1$ denotes the portion of the maximal distance for second layer reception normalized by the maximal distance for first layer reception. As a result, the SNR gap between distances D_1 and D_2 is computed as $\Delta\gamma = \rho^{-\alpha}$.

If we require the first layer reception to include two bits of information, the SNR at the receiver has to be at least equal to that of QPSK, or 4-QAM, reception (see Figure 6). If we utilize a multi-layered CQAM$(f, 2)$ scheme with 16 symbols instead, the additional energy required to guarantee the same first layer reception quality can be calculated as the difference between the dashed red line and the dotted black line in Figure 6. From the same figure, we obtain the SNR gap $\Delta\gamma$ between second layer and first layer reception as the difference between the blue line and the dashed red line. With the above, it is possible to convert this SNR gap into a distance quotient ρ, computed as $\rho = (\Delta\gamma)^{-1/\alpha}$. Thus, we achieve a trade-off of investing an amount of additional energy required for multi-layered CQAM (compared to low-layer QAM) at the gain of a second layer of reception at receivers within a portion ρ of the overall distance. For path loss exponents α ranging between 2 and 4, this trade-off between ρ and additional energy is visualized in Figure 7. For example, at $\alpha = 3$, the energy cost for a full 4-bit reception at 50% of the distance ($\rho = 0.5$) is approximately 2.5 dB.

VII. CONCLUSION

In this paper, we present a novel perspective on the classification and parametrization of practical hierarchical modulation schemes. We introduce generalized Cantor sets and CPAM, a pulse amplitude modulation scheme based on Cantor sets. This scheme is parametrized through an appropriate choice of scaling factors f_i. We show that CPAM classifies arbitrary hierarchical PAM schemes and present a formula for the average
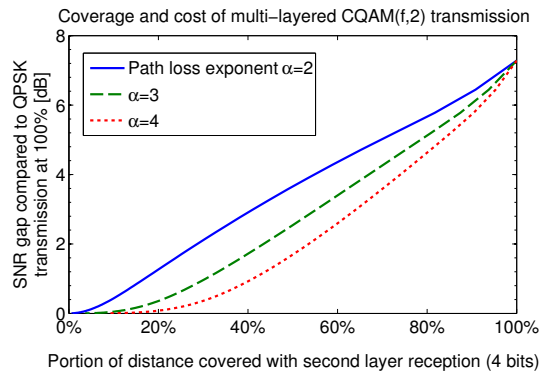
symbol energy, a low-complexity demodulation procedure, as well as a natural description of the boundary points of the decision regions. All of the results can be extended to two dimensional hierarchical QAM schemes by replacing Cantor sets with Cantor dust. Finally, we present exemplary numerical symbol error rates and study the beneficial effects of path loss on hierarchical information transmission.

REFERENCES

[1] T. Cover, "Broadcast channels," *Information Theory, IEEE Transactions on*, vol. 18, no. 1, pp. 2 – 14, Jan. 1972.

[2] A. Calderbank and N. Seshadri, "Multilevel codes for unequal error protection," in *Information Theory, 1993. Proceedings. 1993 IEEE International Symposium on*, Jan. 1993, p. 183.

[3] L.-F. Wei, "Coded modulation with unequal error protection," *Communications, IEEE Transactions on*, vol. 41, no. 10, pp. 1439 –1449, Oct. 1993.

[4] Q. Zhang, W. Zhu, and Y.-Q. Zhang, "Channel-adaptive resource allocation for scalable video transmission over 3G wireless network," *Circuits and Systems for Video Technology, IEEE Transactions on*, vol. 14, no. 8, pp. 1049 – 1063, Aug. 2004.

[5] B. Barmada, M. Ghandi, E. Jones, and M. Ghanbari, "Prioritized transmission of data partitioned H.264 video with hierarchical QAM," *Signal Processing Letters, IEEE*, vol. 12, no. 8, pp. 577 – 580, Aug. 2005.

[6] T. Schierl, T. Stockhammer, and T. Wiegand, "Mobile video transmission using scalable video coding," *Circuits and Systems for Video Technology, IEEE Transactions on*, vol. 17, no. 9, pp. 1204 –1217, Sept. 2007.

[7] P. Vitthaladevuni and M.-S. Alouini, "A recursive algorithm for the exact BER computation of generalized hierarchical QAM constellations," *Information Theory, IEEE Transactions on*, vol. 49, no. 1, pp. 297 – 307, Jan. 2003.

[8] ——, "A closed-form expression for the exact BER of generalized PAM and QAM constellations," *Communications, IEEE Transactions on*, vol. 52, no. 5, pp. 698 – 700, May 2004.

[9] H. Jiang and P. Wilford, "A hierarchical modulation for upgrading digital broadcast systems," *Broadcasting, IEEE Transactions on*, vol. 51, no. 2, pp. 223 – 229, June 2005.

[10] N. Souto, F. Cercas, R. Dinis, and J. Silva, "On the BER performance of hierarchical M-QAM constellations with diversity and imperfect channel estimation," *Communications, IEEE Transactions on*, vol. 55, no. 10, pp. 1852 –1856, Oct. 2007.

[11] A. Goldsmith, *Wireless Communications*. New York, NY, USA: Cambridge University Press, 2005.

[12] L. Steen and J. Seebach, *Counterexamples in topology*, ser. Dover books on mathematics. Dover Publications, 1995.

[13] W. Weber III, "Differential encoding for multiple amplitude and phase shift keying systems," *Communications, IEEE Transactions on*, vol. 26, no. 3, pp. 385 – 391, Mar. 1978.