Optimized Unequal Error Protection Using Multiplexed Hierarchical Modulation

Seok-Ho Chang, *Member, IEEE,* Minjoong Rim, *Member, IEEE,*
Pamela C. Cosman, *Fellow, IEEE,* and Laurence B. Milstein, *Fellow, IEEE*

Abstract

With progressive image or scalable video encoders, as more bits are received, the source image or video can be reconstructed with progressively better quality at the receiver. These progressive codes have gradual differences of importance in their bitstreams, which necessitates multiple levels of unequal error protection (UEP). One practical method of achieving UEP is based on a constellation of nonuniformly spaced signal points, or hierarchical constellations. However, hierarchical modulation can achieve only a limited number of UEP levels for a given constellation size. Though hierarchical modulation has been intensively studied for digital broadcasting or multimedia transmission, most work has considered only two layered source coding, and methods of achieving a large number of levels of UEP for progressive transmission have rarely been studied. In this paper, we propose a multilevel UEP system using multiplexed hierarchical quadrature amplitude modulation (QAM) for progressive transmission over mobile radio channels. We show that multiple levels of UEP are achieved by the proposed method. When the BER is dominated by the minimum Euclidian distance, we derive an optimal multiplexing approach which minimizes both the average and peak powers. We next propose an asymmetric hierarchical QAM which reduces the peak-to-average power ratio (PAPR) of the proposed UEP system without any performance loss. Numerical results show that the performance of progressive transmission over Rayleigh fading channels is significantly enhanced by the proposed multiplexing methods.

S.-H. Chang, P. C. Cosman and L. B. Milstein are with the Department of Electrical and Computer Engineering, University of California, San Diego, La Jolla, CA, 92093–0407, USA (e-mail: s4chang@ucsd.edu; pcosman@ucsd.edu; milstein@ece.ucsd.edu). M. Rim is with the Department of Information and Communication Engineering, Dongguk University, Seoul, Korea (e-mail: minjoong@dongguk.edu).

This work was supported in part by the SETsquared UK/US Programme for Applied Collaborative Research and by the US Army Research Office under MURI, grant number W911NF-04-1-0224.

Index Terms

Cross-layer, hierarchical modulation, multimedia communications, progressive image, scalable video, unequal error protection, wireless video.

I. Introduction

When a communication system transmits messages over mobile radio channels, they are subject to errors, in part because mobile channels typically exhibit time-variant channel-quality fluctuations. For two-way communication links, these effects can be mitigated using adaptive methods [1]–[3]. However, the adaptive schemes require a reliable feedback link from the receiver to the transmitter. Moreover, for a one-way broadcast system, those schemes are not appropriate because of the nature of broadcasting. When adaptive schemes cannot be used, the way to ensure communications is to classify the data into multiple classes with unequal error protection (UEP). The most important class should be recovered by the receiver even under poor receiving conditions. Hence, strong error protection is used for the important data all of the time, even though sometimes there is no need for it. Less important data is always protected less even though sometimes it cannot be recovered successfully.

Theoretical investigation of efficient communication from a single source to multiple receivers established the fundamental idea that optimal broadcast transmission could be achieved by a superposition or hierarchical transmission scheme [4]–[6]. Since the theoretical and conceptual basis for UEP was initiated by Cover [4], much of the work has shown that one practical method of achieving UEP is based on a constellation of nonuniformly spaced signal points [7]–[10], which is called a hierarchical, embedded, or multi-resolution constellation. In this constellation, more important bits in a symbol have larger minimum Euclidian distance than less important bits. Hierarchical constellations were previously considered in [11], and intensively studied for digital broadcasting systems [7][9][10]. Ramchandran *et al.* [7] designed an overall multiresolution digital HDTV broadcast system using hierarchical modulation under a joint source-channel coding (JSCC) framework. Calderbank and Seshadri [9] considered the use of hierarchical quadrature amplitude modulation (QAM) as the adaptive constellations for digital video broadcasting. Moreover, the Digital Video Broadcasting (DVB-T) standard [12], which is now commercially available, incorporated hierarchical QAM for layered video data transmission, since it provides enhanced system-level capacity and coverage in a wireless environment

[13][14]. Pursley and Shea [15][16] also proposed communication systems based on hierarchical modulation which support multimedia transmission by simultaneously delivering different types of traffic, each with its own required quality of service.

Another well known and obvious method to achieve UEP is based on channel coding: more powerful error-correction coding is applied to a more important data class. Block codes for providing UEP were studied by Masnick and Wolf [17], and Suda and Miki [18]. The use of rate-compatible punctured convolutional (RCPC) codes to achieve UEP was suggested by Cox *et al.* [19]. These UEP methods based on error-correction coding have been widely used for layered video or image transmission [20]–[23]. Sometimes, UEP approaches based on hierarchical modulation and error-correction coding were jointly employed in a system [8][9][12][15][23]. For example, in the DVB-T standard [12], two different layers of video data are channel encoded with corresponding coding rates, and then they are mapped to hierarchical 16 or 64 QAM constellation. Pei and Modestino [23] showed that when error-correction coding approach for UEP and hierarchical modulation are jointly used, more efficient and flexible UEP is achieved. Hierarchical modulation has other desirable properties in addition to performance considerations. The amount of UEP can be adjusted in a continuous manner by modifying the spacing between signal points of the constellation [8], and different levels of protection are achieved without an increase in bandwidth compared to channel coding [24].

Progressive image or scalable video encoders [25]–[30], which are expected to have more prominence in the future, employ a mode of transmission such that as more bits are received, the source can be reconstructed with better quality at the receiver. In other words, the decoder can use each additional received bit to improve the quality of the previously reconstructed images. Since these progressive transmissions have gradual differences of importance in their bitstreams, multiple levels of error protection are required. However, unlike channel coding for UEP, hierarchical modulation can achieve only a limited number of UEP levels for a given constellation size. For example, hierarchical 16 QAM provides two levels of UEP, and hierarchical 64 QAM yields at most three levels [31]. In the DVB-T standard, video data encoded by MPEG-2 consists of two different layers, and thus the use of hierarchical 16 or 64 QAM meets the required number of UEP levels. However, if scalable video is to be incorporated in a digital video broadcasting system, hierarchical 16 or 64 QAM may not meet the system needs. Most of the work about hierarchical modulation up to now has been restricted to consideration of two layered source coding, and methods of achieving a large number of levels of UEP for progressive mode of transmission have rarely been studied.

In this paper, we propose a multilevel UEP system using multiplexed hierarchical modulation for progressive transmission over mobile radio channels. We propose a way of multiplexing hierarchical QAM constellations, and show that arbitrarily large number of UEP levels are achieved by the proposed method. These results are presented in Section II. When the BER is dominated by the minimum Euclidian distance, we derive an an optimal multiplexing approach which minimizes both the average and peak powers, which is presented in Section III. While the suggested methods achieve multilevel UEP, the PAPR typically will be increased when constellations having distinct minimum distances are time-multiplexed. To mitigate this effect, an asymmetric hierarchical QAM constellation, which reduces the PAPR without performance loss, is designed in Section IV. In Section V, we consider the case where multiplexed constellations need to have constant power, either due to the limited capability of a power amplifier, or for the ease of cochannel interference control. In Section VI, the performance of the suggested UEP system for the transmission of progressive images is analyzed in terms of the expected distortion, and Section VII presents numerical results of performance analysis.

II. Multilevel UEP Based on Multiplexing Hierarchical QAM Constellations

A. *Hierarchical 16 QAM Constellation*

First, we analyze hierarchical 16 QAM as a special case. Fig. 1 shows a hierarchical 16 QAM constellation with Gray coded bit mapping [12]. The 16 signal points are divided into four clusters and each cluster consists of four signal points. The two most significant bits (MSBs), i_1 and q_1, determine one of the four clusters, and their minimum Euclidian distance is d_M. The two least significant bits (LSBs), i_2 and q_2, determine which of the four signal points within the cluster is chosen, and their minimum Euclidian distance is d_L. The distance ratio $\alpha = d_M/d_L$ (> 1) determines how much more the MSBs are protected against errors than are the LSBs. Hierarchical 16 QAM has one embedded QPSK subconstellation consisting of four clusters, and thus is denoted by 4/16 QAM.

We consider multiplexing N hierarchical 16 QAM constellations, all of which have distinct minimum distances. The average power per symbol of all the multiplexed constellations, S_{avg},

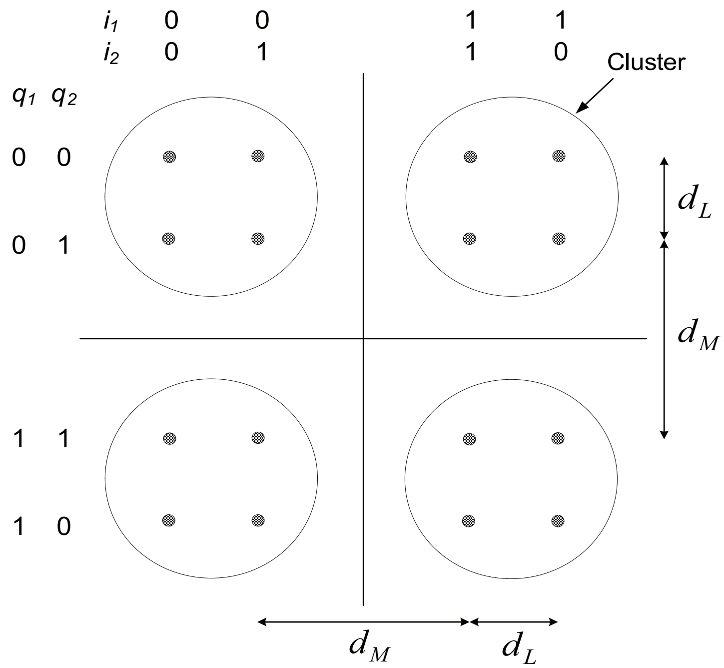

Fig. 1. Hierarchical 16 QAM constellation.

is given by

$$S_{avg} = \frac{1}{N}\sum_{i=1}^{N} S_{avg,i} \tag{1}$$

where $S_{avg,i}$ is the average power per symbol of constellation i. For hierarchical 16 QAM, $S_{avg,i}$ is given by

$$S_{avg,i} = \left(\frac{d_{M,i}}{2}\right)^2 + \left(\frac{d_{M,i}}{2} + d_{L,i}\right)^2 = \frac{d_{M,i}^2}{2} + d_{M,i}d_{L,i} + d_{L,i}^2 \tag{2}$$

where $d_{M,i}$ and $d_{L,i}$ are minimum distances for the MSBs and LSBs of constellation i, respectively. The BERs of the MSBs and LSBs of hierarchical 16 QAM constellation i, denoted by $P_{M,i}$ and $P_{L,i}$, respectively, are given by [31]

$$\begin{aligned} P_{M,i} &= \frac{1}{2}Q\left(\frac{d_{M,i}}{2}\sqrt{\frac{2\gamma_s}{S_{avg}}}\right) + \frac{1}{2}Q\left(\left(\frac{d_{M,i}}{2} + d_{L,i}\right)\sqrt{\frac{2\gamma_s}{S_{avg}}}\right) \\ P_{L,i} &= Q\left(\frac{d_{L,i}}{2}\sqrt{\frac{2\gamma_s}{S_{avg}}}\right) + \frac{1}{2}Q\left(\left(d_{M,i} + \frac{d_{L,i}}{2}\right)\sqrt{\frac{2\gamma_s}{S_{avg}}}\right) - \frac{1}{2}Q\left(\left(d_{M,i} + \frac{3d_{L,i}}{2}\right)\sqrt{\frac{2\gamma_s}{S_{avg}}}\right) \end{aligned} \tag{3}$$

where S_{avg} is given by (1) and (2), γ_s is the signal-to-noise ratio (SNR) per symbol, and $Q(x) = 1/\sqrt{2\pi}\int_x^{\infty} e^{-y^2/2}dy$.

The following theorem states that $2N$ levels of UEP can be achieved by multiplexing N hierarchical 16 QAM constellations.

Theorem 1: For N hierarchical 16 QAM constellations, $P_{M,i}$ and $P_{L,i}$, given by (3), satisfy

$$P_{M,1} < P_{M,2} < \cdots < P_{M,N} < P_{L,1} < P_{L,2} < \cdots < P_{L,N} \tag{4}$$

for all SNR if

$$d_{M,1} > d_{M,2} > \cdots > d_{M,N} > d_{L,1} > d_{L,2} > \cdots > d_{L,N}. \tag{5}$$

Proof: We will first show that, for $1 \le i, j \le N$,

$$P_{M,i} < P_{L,j} \quad \text{if} \quad d_{M,i} > d_{L,j}. \tag{6}$$

Since $Q(x)$ is a monotonically decreasing function, from (3), we have

$$P_{M,i} < \frac{1}{2}Q\left(\frac{d_{M,i}}{2}\sqrt{\frac{2\gamma_s}{S_{avg}}}\right) + \frac{1}{2}Q\left(\frac{d_{M,i}}{2}\sqrt{\frac{2\gamma_s}{S_{avg}}}\right) = Q\left(\frac{d_{M,i}}{2}\sqrt{\frac{2\gamma_s}{S_{avg}}}\right). \tag{7}$$

If $d_{M,i} > d_{L,j}$, from (3) and (7), we have

$$P_{M,i} < Q\left(\frac{d_{L,j}}{2}\sqrt{\frac{2\gamma_s}{S_{avg}}}\right) < P_{L,j}. \tag{8}$$

We next show that, for $d_{M,1} > d_{M,2} > \cdots > d_{M,N}$ and $d_{L,1} > d_{L,2} > \cdots > d_{L,N}$,

$$P_{M,1} < P_{M,2} < \cdots < P_{M,N}. \tag{9}$$

Consider two constellations i and $i+1$ among N hierarchical constellations ($1 \le i \le N-1$). From (3), we have $P_{M,i} < P_{M,i+1}$ if $d_{M,i} > d_{M,i+1}$ and $d_{L,i} > d_{L,i+1}$.

Lastly, we show that for $d_{M,1} > d_{M,2} > \cdots > d_{M,N}$ and $d_{L,1} > d_{L,2} > \cdots > d_{L,N}$,

$$P_{L,1} < P_{L,2} < \cdots < P_{L,N}. \tag{10}$$

We define a function $f(x,y)$ as

$$f(x,y) = Q\left(\frac{y}{2}\right) + \frac{1}{2}Q\left(x + \frac{y}{2}\right) - \frac{1}{2}Q\left(x + \frac{3y}{2}\right). \tag{11}$$

$f(x,y)$ is a monotonically decreasing function of $x > 0$ and $y > 0$, since

$$\begin{aligned} \frac{\partial f(x,y)}{\partial x} &= \frac{-1}{2\sqrt{2\pi}}\left[e^{-\frac{1}{2}\left(x+\frac{y}{2}\right)^2} - e^{-\frac{1}{2}\left(x+\frac{3y}{2}\right)^2}\right] < 0, \quad \text{and} \\ \frac{\partial f(x,y)}{\partial y} &= \frac{-1}{2\sqrt{2\pi}}\left[e^{-\frac{1}{2}\left(\frac{y}{2}\right)^2} - e^{-\frac{1}{2}\left(x+\frac{3y}{2}\right)^2} + \frac{1}{2}\left\{e^{-\frac{1}{2}\left(x+\frac{y}{2}\right)^2} - e^{-\frac{1}{2}\left(x+\frac{3y}{2}\right)^2}\right\}\right] < 0. \end{aligned} \tag{12}$$

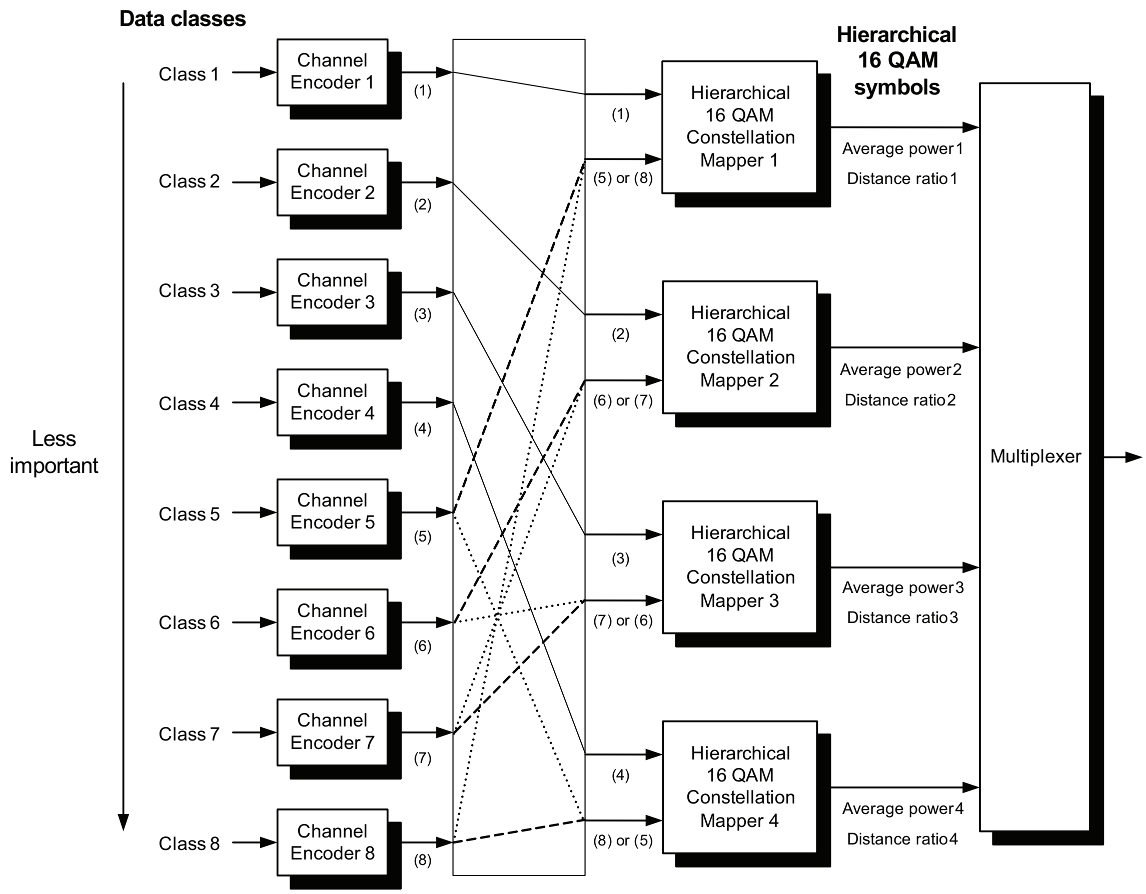

Fig. 2. The multilevel UEP system using multiplexed hierarchical 16 QAM constellations. (a) The system based on Corollary 2: dashed lines (b) The system based on Corollary 9: dotted lines (note that solid lines are for both Corollaries 2 and 9).

From (3) and (11), it is seen that $P_{L,i} = f\left(d_{M,i}\sqrt{2\gamma_s/S_{avg}},\ d_{L,i}\sqrt{2\gamma_s/S_{avg}}\right)$. Hence, from (12), we have

$$P_{L,i} < P_{L,i+1} \quad \text{if} \quad d_{M,i} > d_{M,i+1} \;\; \text{and} \;\; d_{L,i} > d_{L,i+1}. \tag{13}$$

Finally, (4) and (5) are derived from (6), (9) and (10).

□

Theorem 1 tells us that $2N$ levels of UEP are achieved by multiplexing N hierarchical 16 QAM constellations having the minimum distances satisfying (5).

Corollary 2: Suppose that there are $2N$ unequally important data classes to be transmitted, and class i is more important than class $i+1$ for $1 \le i \le 2N-1$. Let P_i denote the BER of data class i. Then,

$$P_1 < P_2 < \cdots < P_{2N} \tag{14}$$

is satisfied for all SNR if the following conditions hold:

i) Class i and class $N+i$ are mapped to the MSBs and LSBs of constellation i, respectively, $(1 \leq i \leq N)$.

ii) The minimum Euclidian distances of the constellations satisfy (5).

Proof: If i) is satisfied, P_i is given by

$$P_i = P_{M,i} \quad \text{and} \quad P_{N+i} = P_{L,i} \quad (1 \leq i \leq N). \tag{15}$$

If ii) is satisfied, we have $P_{M,1} < P_{M,2} < \cdots < P_{M,N} < P_{L,1} < P_{L,2} < \cdots < P_{L,N}$ from Theorem 1.

□

Fig. 2 (a) depicts the multilevel UEP system using multiplexed hierarchical 16 QAM constellations based on Corollary 2 for eight data classes ($N = 4$).

B. Hierarchical 2^{2K} $(K \geq 3)$ *QAM Constellation*

Next, we consider multiplexing hierarchical 2^{2K} $(K \geq 3)$ QAM constellations. As an example, Fig. 3 depicts a hierarchical 64 QAM constellation ($K = 3$). The two MSBs i_1 and q_1 determine the quadrant of the first cluster, and their minimum Euclidian distance is d_{M1}. The second two MSBs i_2 and q_2 determine the quadrant within the first cluster, and their minimum distance is d_{M2}. Lastly, the third two MSBs (or LSBs) i_3 and q_3 determine the symbol within the second cluster, and their minimum distance is d_{M3}. Hierarchical 64 QAM has two embedded subconstellations, and thus is denoted by 4/16/64 QAM. The hierarchical 64 QAM operates as QPSK when channel conditions are poor, and it operates as 16 or 64 QAM when channel quality gets better. The BER of hierarchical 2^{2K} QAM, P_{M_n}, is given by a recursive expression in [31].

In the following lemma, the BERs of hierarchical 2^{2K} QAM are derived under some assumption based on the fact that for hierarchical constellations, minimum distance for more important bits is greater than that for less important bits.

Lemma 3: Let d_{M_n} denote the minimum distance for the nth MSBs $(1 \leq n \leq K)$. Note that the distance ratio of the hierarchical constellation, $d_{M_{n-1}}/d_{M_n}$, is greater than unity ($2 \leq n \leq K$).

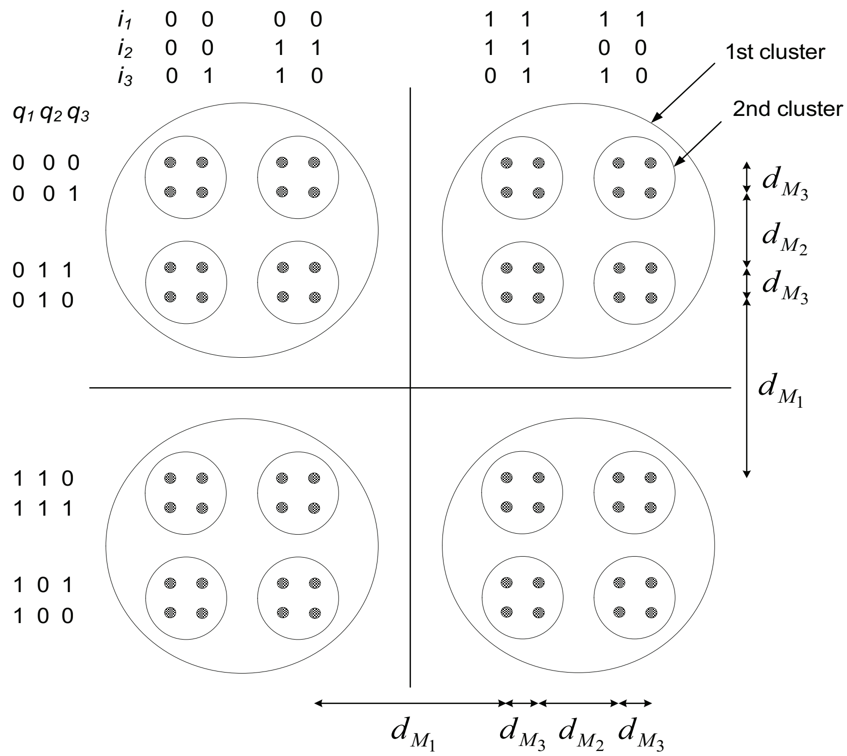

Fig. 3. Hierarchical 64 QAM constellation.

If the SNR of interest for the nth MSBs is sufficiently large so that the probability of the noise exceeding the Euclidian distance of $d_{M_{n-1}} + \frac{1}{2}d_{M_n}$ is insignificant compared to that of the noise exceeding $\frac{1}{2}d_{M_n}$, the BER of the nth MSBs ($2 \leq n \leq K$), P_{M_n}, becomes

$$P_{M_n}^{app} = \begin{cases} \sum_{p=0}^{2^{K-n}-1} \frac{1}{2^{K-n}} Q\left(\left(\frac{d_{M_n}}{2} + \sum_{q=n+1}^{K} \left\lfloor \frac{p+2^{K-q}}{2^{K-q+1}} \right\rfloor d_{M_q}\right)\sqrt{\frac{2\gamma_s}{S_{avg}}}\right), \\ \text{for } 2 \leq n \leq K-1 \\ Q\left(\frac{d_{M_K}}{2}\sqrt{\frac{2\gamma_s}{S_{avg}}}\right) + \frac{1}{2}Q\left(\left(d_{M_{K-1}} + \frac{d_{M_K}}{2}\right)\sqrt{\frac{2\gamma_s}{S_{avg}}}\right), \\ \text{for } n = K \end{cases} \quad (16)$$

where $\lfloor x \rfloor$ denotes the largest integer less than or equal to x, and $S_{avg} = \sum_{u=1}^{K}\sum_{v=u}^{K} \mu_{uv} d_{M_u} d_{M_v}$ is the average power of a hierarchical 2^{2K} QAM, where the μ_{uv} are constants. Note that for the MSBs (i.e., $n = 1$), the top line of (16) is the exact BER expression when n is set to unity (i.e., $P_{M_1}^{app} = P_{M_1}$).

Proof: See Appendix A.

□

$P^{app}_{M_n}$ is numerically evaluated for hierarchical 64 and 256 QAM in Appendix B as an example. For both constellations, $P^{app}_{M_n}$ $(2 \leq n \leq K)$ is shown to be close to the exact BER within 0.001 dB for BER ≤ 0.1 even at the lower bound of the distance ratio (i.e., $d_{M_{n-1}}/d_{M_n} = 1$). Note that for reference, the distance ratio of hierarchial 16 and 64 QAM in the DVB-T standard [12] is 2 or 4.

For N multiplexed hierarchical 2^{2K} QAM constellations, the average power per symbol of constellation i is given by

$$S_{avg,i} = \sum_{u=1}^{K}\sum_{v=u}^{K} \mu_{uv} d_{M_u,i} d_{M_v,i} \tag{17}$$

where $d_{M_n,i}$ $(1 \leq n \leq K)$ is the minimum distance for the nth MSBs of constellation i $(1 \leq i \leq N)$, and the μ_{uv} are constants. When the condition of Lemma 3 is satisfied, from (1), (16) and (17), the BER of the nth MSBs $(2 \leq n \leq K)$ of a hierarchical 2^{2K} QAM constellation i, $P_{M_n,i}$, becomes

$$P^{app}_{M_n,i} = \begin{cases} \sum_{p=0}^{2^{K-n}-1} \frac{1}{2^{K-n}} Q\left(\left(\frac{d_{M_n,i}}{2} + \sum_{q=n+1}^{K} \left\lfloor \frac{p+2^{K-q}}{2^{K-q+1}} \right\rfloor d_{M_q,i}\right)\sqrt{\frac{2\gamma_s}{S_{avg}}}\right), \\ \text{for} \quad 2 \leq n \leq K-1 \\ Q\left(\frac{d_{M_k,i}}{2}\sqrt{\frac{2\gamma_s}{S_{avg}}}\right) + \frac{1}{2} Q\left(\left(d_{M_{K-1},i} + \frac{d_{M_K,i}}{2}\right)\sqrt{\frac{2\gamma_s}{S_{avg}}}\right), \\ \text{for} \quad n = K. \end{cases} \tag{18}$$

Note that the top line of (18) is the exact BER expression when n is set to unity (i.e., $P^{app}_{M_1,i} = P_{M_1,i}$).

Theorem 4: For N hierarchical 2^{2K} QAM constellations, $P^{app}_{M_n,i}$, given by (18), satisfy

$$P^{app}_{M_1,1} < \cdots < P^{app}_{M_1,N} < P^{app}_{M_2,1} < \cdots < P^{app}_{M_2,N} < \cdots < P^{app}_{M_K,1} < \cdots < P^{app}_{M_K,N} \tag{19}$$

$$\text{if} \quad d_{M_1,1} > \cdots > d_{M_1,N} > d_{M_2,1} > \cdots > d_{M_2,N} > \cdots > d_{M_K,1} > \cdots > d_{M_K,N}. \tag{20}$$

Proof: We will first show that, for $1 \leq i, j \leq N$,

$$\begin{aligned} &P^{app}_{M_1,i} < P^{app}_{M_2,j},\ P^{app}_{M_2,i} < P^{app}_{M_3,j}, \cdots, P^{app}_{M_{K-1},i} < P^{app}_{M_K,j} \\ &\text{if} \quad d_{M_1,i} > d_{M_2,j},\ d_{M_2,i} > d_{M_3,j}, \cdots, d_{M_{K-1},i} > d_{M_K,j}. \end{aligned} \tag{21}$$

From (18), $P^{app}_{M_n,i}$ $(1 \leq n \leq K-2)$ can be expressed as

$$P^{app}_{M_n,i} = \sum_{r=0}^{2^{K-n-1}-1} \frac{1}{2^{K-n}} Q\left(\left(\frac{d_{M_n,i}}{2} + \sum_{q=n+1}^{K} \left\lfloor \frac{2r+2^{K-q}}{2^{K-q+1}} \right\rfloor d_{M_q,i}\right)\sqrt{\frac{2\gamma_s}{S_{avg}}}\right)$$
$$+ \sum_{r=0}^{2^{K-n-1}-1} \frac{1}{2^{K-n}} Q\left(\left(\frac{d_{M_n,i}}{2} + \sum_{q=n+1}^{K} \left\lfloor \frac{2r+1+2^{K-q}}{2^{K-q+1}} \right\rfloor d_{M_q,i}\right)\sqrt{\frac{2\gamma_s}{S_{avg}}}\right). \quad (22)$$

Eq. (22) can be rewritten as

$$P^{app}_{M_n,i} = \sum_{r=0}^{2^{K-n-1}-1} \frac{1}{2^{K-n}} Q\left(\left(\frac{d_{M_n,i}}{2} + \sum_{q=n+1}^{K} \left\lfloor \frac{r+2^{K-q-1}}{2^{K-q}} \right\rfloor d_{M_q,i}\right)\sqrt{\frac{2\gamma_s}{S_{avg}}}\right)$$
$$+ \sum_{r=0}^{2^{K-n-1}-1} \frac{1}{2^{K-n}} Q\left(\left(\frac{d_{M_n,i}}{2} + \sum_{q=n+1}^{K} \left\lfloor \frac{r+2^{-1}+2^{K-q-1}}{2^{K-q}} \right\rfloor d_{M_q,i}\right)\sqrt{\frac{2\gamma_s}{S_{avg}}}\right). \quad (23)$$

From (23), since $r+2^{K-q-1}$ and 2^{K-q} are integers for $q \leq K-1$, we have

$$\left\lfloor \frac{r+2^{-1}+2^{K-q-1}}{2^{K-q}} \right\rfloor = \left\lfloor \frac{r+2^{K-q-1}}{2^{K-q}} \right\rfloor \quad \text{for } q \leq K-1. \quad (24)$$

From (23), for $q = K$, we have

$$\left\lfloor \frac{r+2^{K-q-1}}{2^{K-q}} \right\rfloor = r \quad \text{and} \quad \left\lfloor \frac{r+2^{-1}+2^{K-q-1}}{2^{K-q}} \right\rfloor = r+1. \quad (25)$$

From (24) and (25), (23) can be rewritten as

$$P^{app}_{M_n,i} = \sum_{r=0}^{2^{K-n-1}-1} \frac{1}{2^{K-n}} Q\left(\left(\frac{d_{M_n,i}}{2} + \sum_{q=n+1}^{K-1} \left\lfloor \frac{r+2^{K-q-1}}{2^{K-q}} \right\rfloor d_{M_q,i} + r d_{M_K,i}\right)\sqrt{\frac{2\gamma_s}{S_{avg}}}\right)$$
$$+ \sum_{r=0}^{2^{K-n-1}-1} \frac{1}{2^{K-n}} Q\left(\left(\frac{d_{M_n,i}}{2} + \sum_{q=n+1}^{K-1} \left\lfloor \frac{r+2^{K-q-1}}{2^{K-q}} \right\rfloor d_{M_q,i} + (r+1) d_{M_K,i}\right)\sqrt{\frac{2\gamma_s}{S_{avg}}}\right). \quad (26)$$

Setting $t = q+1$, $P^{app}_{M_n,i}$ $(1 \leq n \leq K-2)$, given by (26), can be expressed as

$$P^{app}_{M_n,i} = \sum_{r=0}^{2^{K-n-1}-1} \frac{1}{2^{K-n}} Q\left(\left(\frac{d_{M_n,i}}{2} + \sum_{t=n+2}^{K} \left\lfloor \frac{r+2^{K-t}}{2^{K-t+1}} \right\rfloor d_{M_{t-1},i} + r d_{M_K,i}\right)\sqrt{\frac{2\gamma_s}{S_{avg}}}\right)$$
$$+ \sum_{r=0}^{2^{K-n-1}-1} \frac{1}{2^{K-n}} Q\left(\left(\frac{d_{M_n,i}}{2} + \sum_{t=n+2}^{K} \left\lfloor \frac{r+2^{K-t}}{2^{K-t+1}} \right\rfloor d_{M_{t-1},i} + (r+1) d_{M_K,i}\right)\sqrt{\frac{2\gamma_s}{S_{avg}}}\right). \quad (27)$$

From (18), $P^{app}_{M_{n+1},j}$ $(1 \leq n \leq K-2)$ can be rewritten as

$$P^{app}_{M_{n+1},j} = \sum_{p=0}^{2^{K-n-1}-1} \frac{1}{2^{K-n}} Q\left(\left(\frac{d_{M_{n+1},j}}{2} + \sum_{q=n+2}^{K} \left\lfloor \frac{p+2^{K-q}}{2^{K-q+1}} \right\rfloor d_{M_q,j}\right)\sqrt{\frac{2\gamma_s}{S_{avg}}}\right) + \sum_{p=0}^{2^{K-n-1}-1} \frac{1}{2^{K-n}} Q\left(\left(\frac{d_{M_{n+1},j}}{2} + \sum_{q=n+2}^{K} \left\lfloor \frac{p+2^{K-q}}{2^{K-q+1}} \right\rfloor d_{M_q,j}\right)\sqrt{\frac{2\gamma_s}{S_{avg}}}\right). \quad (28)$$

From (27) and (28), for $1 \leq n \leq K-2$, we have

$$P^{app}_{M_n,i} < P^{app}_{M_{n+1},j} \quad \text{if} \quad d_{M_n,i} > d_{M_{n+1},j},\ d_{M_{n+1},i} > d_{M_{n+2},j}, \cdots, d_{M_{K-1},i} > d_{M_K,j}. \quad (29)$$

From (18), $P^{app}_{M_{K-1},i}$ is given by

$$P^{app}_{M_{K-1},i} = \frac{1}{2} Q\left(\frac{d_{M_{K-1},i}}{2}\sqrt{\frac{2\gamma_s}{S_{avg}}}\right) + \frac{1}{2} Q\left(\left(\frac{d_{M_{K-1},i}}{2} + d_{M_K,i}\right)\sqrt{\frac{2\gamma_s}{S_{avg}}}\right). \quad (30)$$

From (18) and (30), we have

$$P^{app}_{M_{K-1},i} < P^{app}_{M_K,j} \quad \text{if} \quad d_{M_{K-1},i} > d_{M_K,j}. \quad (31)$$

From (29) and (31), (21) is derived.

We next show that

$$P^{app}_{M_1,1} < \cdots < P^{app}_{M_1,N},\ P^{app}_{M_2,1} < \cdots < P^{app}_{M_2,N}, \cdots, P^{app}_{M_K,1} < \cdots < P^{app}_{M_K,N}$$
$$\text{if} \quad d_{M_1,1} > \cdots > d_{M_1,N},\ d_{M_2,1} > \cdots > d_{M_2,N}, \cdots, d_{M_K,1} > \cdots < d_{M_K,N}. \quad (32)$$

We define a function $f(x_n, x_{n+1}, \cdots, x_K)$ as

$$f(x_n, x_{n+1}, \cdots, x_K) = \sum_{p=0}^{2^{K-n}-1} \frac{1}{2^{K-n}} Q\left(\left(\frac{x_n}{2} + \sum_{q=n+1}^{K} \left\lfloor \frac{p+2^{K-q}}{2^{K-q+1}} \right\rfloor x_q\right)\right). \quad (33)$$

The $f(x_n, x_{n+1}, \cdots, x_K)$ is a monotonically decreasing function of $x_n > 0, x_{n+1} > 0, \cdots, x_K > 0$, since

$$\frac{\partial f(x_n, x_{n+1}, \cdots, x_K)}{\partial x_n} = \frac{-1}{2\sqrt{2\pi}} \sum_{p=0}^{2^{K-n}-1} \frac{1}{2^{K-n}} e^{-\frac{1}{2}\left(\frac{x_n}{2} + \sum_{q=n+1}^{K} \left\lfloor \frac{p+2^{K-q}}{2^{K-q+1}} \right\rfloor x_q\right)^2} < 0, \quad \text{and}$$
$$\frac{\partial f(x_n, x_{n+1}, \cdots, x_K)}{\partial x_{n+m}} = \frac{-1}{\sqrt{2\pi}} \sum_{p=0}^{2^{K-n}-1} \frac{1}{2^{K-n}} e^{-\frac{1}{2}\left(\frac{x_n}{2} + \sum_{q=n+1}^{K} \left\lfloor \frac{p+2^{K-q}}{2^{K-q+1}} \right\rfloor x_q\right)^2} \left\lfloor \frac{p+2^{K-n-m}}{2^{K-n-m+1}} \right\rfloor < 0 \quad (34)$$

for $m = 1, \cdots, K-n$ (i.e., for $x_{n+1}, \cdots, x_K$). From (18) and (33), it is seen that for $1 \leq n \leq K-1$,

$$P^{app}_{M_n,i} = f\left(d_{M_n,i}\sqrt{\frac{2\gamma_s}{S_{avg}}},\ d_{M_{n+1},i}\sqrt{\frac{2\gamma_s}{S_{avg}}}, \cdots, d_{M_K,i}\sqrt{\frac{2\gamma_s}{S_{avg}}}\right). \quad (35)$$

From (34) and (35), for $1 \leq n \leq K-1$, we have

$$P^{app}_{M_n,i} < P^{app}_{M_n,i+1} \ \text{ if } \ d_{M_n,i} > d_{M_n,i+1},\ d_{M_{n+1},i} > d_{M_{n+1},i+1}, \cdots, d_{M_K,i} > d_{M_K,i+1}. \quad (36)$$

From (18), for $n = K$, we have

$$P^{app}_{M_K,i} < P^{app}_{M_K,i+1} \ \text{ if } \ d_{M_{K-1},i} > d_{M_{K-1},i+1} \quad \text{and} \quad d_{M_K,i} > d_{M_K,i+1}. \quad (37)$$

From (36) and (37), the following is derived.

$$\begin{aligned} &P^{app}_{M_1,i} < P^{app}_{M_1,i+1},\ P^{app}_{M_2,i} < P^{app}_{M_2,i+1}, \cdots, P^{app}_{M_K,i} < P^{app}_{M_K,i+1} \\ &\text{if} \quad d_{M_1,i} > d_{M_1,i+1},\ d_{M_2,i} > d_{M_2,i+1}, \cdots, d_{M_K,i} > d_{M_K,i+1}. \end{aligned} \quad (38)$$

With $i = 1, \cdots, N-1$, (38) leads to (32). Finally, from (21) and (32), (19) and (20) are derived. □

Theorem 4 tells us that, by multiplexing N hierarchical 2^{2K} $(K \geq 3)$ QAM constellations having the minimum distances satisfying (20), KN levels of UEP are achieved under the assumption that the SNR of interest for the nth MSBs $(2 \leq n \leq K)$ is reasonably large so that the condition of Lemma 3 is satisfied. We note that there are counter examples showing that KN levels of UEP is not achieved for a very low SNR, even when the minimum distances satisfy (20).

III. Optimal Multiplexing of Hierarchical QAM Constellations for High SNR

In this section, we define high SNR as an SNR which is sufficiently large so that the BER is dominated by the Q-function term having the minimum Euclidian distance.

A. Hierarchical $2^{2J}/2^{2K}$ $(K > J \geq 1)$ QAM Constellation

Hierarchical $2^{2J}/2^{2K}$ QAM refers to a specific kind of hierarchical constellations which provide two levels of UEP. Typical examples are hierarchical 4/16 QAM (i.e., hierarchical 16 QAM) and 4/64 QAM which are employed in DVB-T standard. Similar to Section II, we first analyze a hierarchical 16 QAM as a simple example. For high SNR, from (3), the BERs of a hierarchical 16 QAM constellation i $(1 \leq i \leq N)$ are given by

$$P_{M,i} \approx \frac{1}{2} Q\left(\frac{d_{M,i}}{2}\sqrt{\frac{2\gamma_s}{S_{avg}}}\right) \quad \text{and} \quad P_{L,i} \approx Q\left(\frac{d_{L,i}}{2}\sqrt{\frac{2\gamma_s}{S_{avg}}}\right). \tag{39}$$

Theorem 5: Suppose that there are N multiplexed hierarchical 16 QAM constellations, and the minimum distances satisfying (5) are given. Also suppose the given minimum distances can be permuted such that $d_{M,1}, \cdots, d_{M,N}$ for the MSBs can be arbitrarily combined with $d_{L,1}, \cdots, d_{L,N}$ for the LSBs. After the distances are permuted, the resultant minimum distances for the MSBs and LSBs of constellation i, denoted by $\tilde{d}_{M,i}$ and $\tilde{d}_{L,i}$, respectively, can be expressed as

$$\tilde{d}_{M,i} = d_{M,i} \quad \text{and} \quad \tilde{d}_{L,\pi(i)} = d_{L,i} \tag{40}$$

where $\pi(i)$ is the index of the constellation to which $d_{L,i}$ is permuted. Then, with the permuted distances given by (40), the BERs of the data classes satisfy

$$P_1 < P_2 < \cdots < P_{2N} \tag{41}$$

for high SNR if class i and class $N+i$ are mapped to the MSBs of constellation i and the LSBs of constellation $\pi(i)$, respectively $(1 \leq i \leq N)$.

Proof: After distances are permuted, from (39), (40) and the mapping condition below (41), the BERs of data classes are given by

$$P_i \approx \frac{1}{2} Q\left(\frac{d_{M,i}}{2}\sqrt{\frac{2\gamma_s}{S_{avg}}}\right) \quad \text{and} \quad P_{N+i} \approx Q\left(\frac{d_{L,i}}{2}\sqrt{\frac{2\gamma_s}{S_{avg}}}\right) \quad (1 \leq i \leq N). \tag{42}$$

Since $d_{M,N} > d_{L,1}$ from (5), and from (42), we have $P_N < P_{N+1}$. Since $d_{M,i} > d_{M,i+1}$ and $d_{L,i} > d_{L,i+1}$ $(1 \leq i \leq N-1)$ from (5), and from (42), we have

$$P_i < P_{i+1} \quad \text{and} \quad P_{N+i} < P_{N+1+i} \quad (1 \leq i \leq N-1). \tag{43}$$

Since $P_N < P_{N+1}$, and from (43), it follows that $P_1 < \cdots < P_N < P_{N+1} < \cdots < P_{2N}$.

□

In contrast to Theorem 1 and Corollary 2, Theorem 5 tells us that $2N$ levels of UEP are achieved for high SNR even after the minimum distances satisfying (5) are arbitrarily permuted.

Corollary 6: From Theorem 5, when the minimum distances $d_{M,1}$, $\cdots$, $d_{M,N}$ and $d_{L,1}$, $\cdots$, $d_{L,N}$ are permuted for high SNR, the BERs of the data classes, $P_1, \cdots, P_{2N}$, are unchanged.

Proof: From (42), it is seen that P_i $(1 \leq i \leq 2N)$ is not dependent on the choice of $\pi(i)$.

□

Theorem 7: After the distances are permuted as described in Theorem 5, the average power of all the multiplexed hierarchical 16 QAM constellations, S_{avg}, given by

$$S_{avg} = \frac{1}{N}\sum_{i=1}^{N} S_{avg,i} = \frac{1}{N}\sum_{i=1}^{N}\left(\frac{\tilde{d}_{M,i}^2}{2} + \tilde{d}_{M,i}\tilde{d}_{L,i} + \tilde{d}_{L,i}^2\right) \tag{44}$$

is minimized if and only if distances are permuted such that $d_{M,i}$ is combined with $d_{L,N+1-i}$ in the same constellation. That is,

$$\tilde{d}_{M,i} = d_{M,i} \text{ and } \tilde{d}_{L,i} = d_{L,N+1-i} \quad (1 \leq i \leq N). \tag{45}$$

Proof: We will prove the following by induction on the number of hierarchical constellations: For given distances $d_{M,1} > \cdots > d_{M,N}$ and $d_{L,1} > \cdots > d_{L,N}$,

$$f_N^* = \sum_{i=1}^{N}\left(\frac{d_{M,i}^2}{2} + d_{M,i}d_{L,N+1-i} + d_{L,N+1-i}^2\right) \tag{46}$$

is the minimum of $f_N = \sum_{i=1}^{N}\left(\tilde{d}_{M,i}^2/2 + \tilde{d}_{M,i}\tilde{d}_{L,i} + \tilde{d}_{L,i}^2\right)$.

Consider two constellations (i.e., $N = 2$). For given $d_{M,1} > d_{M,2}$ and $d_{L,1} > d_{L,2}$, the distances can be permuted such that $d_{M,1}$ is combined with either $d_{L,1}$ or $d_{L,2}$. The two possible values of f_2 are given by

$$\begin{aligned} f_{2,\sharp 1} &= \frac{d_{M,1}^2}{2} + d_{M,1}d_{L,1} + d_{L,1}^2 + \frac{d_{M,2}^2}{2} + d_{M,2}d_{L,2} + d_{L,2}^2 \\ f_{2,\sharp 2} &= \frac{d_{M,1}^2}{2} + d_{M,1}d_{L,2} + d_{L,2}^2 + \frac{d_{M,2}^2}{2} + d_{M,2}d_{L,1} + d_{L,1}^2. \end{aligned} \tag{47}$$

The difference between $f_{2,\sharp 1}$ and $f_{2,\sharp 2}$ is given by

$$f_{2,\sharp 1} - f_{2,\sharp 2} = (d_{M,1} - d_{M,2})(d_{L,1} - d_{L,2}) > 0 \tag{48}$$

because $d_{M,1} > d_{M,2}$ and $d_{L,1} > d_{L,2}$. From (48), it is seen that $f_{2,\sharp 2}$ is the minimum. For $N = 2$, f_2^* given by (46) is equal to $f_{2,\sharp 2}$.

Suppose that (46) holds when there are l constellations (i.e., $N = l$). In other words, for given $d_{M,1} > \cdots > d_{M,l}$ and $d_{L,1} > \cdots > d_{L,l}$, $f_l^* = \sum_{i=1}^{l}(d_{M,i}^2/2 + d_{M,i}d_{L,l+1-i} + d_{L,l+1-i}^2)$ is the minimum of f_l. Consider $l+1$ constellations (i.e., $N = l+1$). For given $d_{M,1} > \cdots > d_{M,l+1}$ and $d_{L,1} > \cdots > d_{L,l+1}$, we will prove that if f_{l+1} is minimized, $d_{M,1}$ should be combined with $d_{L,l+1}$ in the same constellation by contradicting the following assumption: f_{l+1} is minimized with $d_{M,1}$ and $d_{L,l+1}$ not being combined. By the assumption, $d_{M,1}$ and $d_{L,j}$ (for some j in the range of $1 \leq j < l+1$) are combined in some specific constellation, and $d_{M,k}$ and $d_{L,l+1}$ (for some k in the range of $1 < k \leq l+1$) are combined in another constellation. The corresponding f_{l+1}, denoted by $f_{l+1,\sharp 1}$, is given by

$$\begin{aligned} f_{l+1,\sharp 1} = & \left(\frac{d_{M,1}^2}{2} + d_{M,1}d_{L,j} + d_{L,j}^2\right) + \left(\frac{d_{M,k}^2}{2} + d_{M,k}d_{L,l+1} + d_{L,l+1}^2\right) \\ & + \sum_{\substack{i=2 \\ i \neq k}}^{l+1} \left(\frac{\tilde{d}_{M,i}^2}{2} + \tilde{d}_{M,i}\tilde{d}_{L,i} + \tilde{d}_{L,i}^2\right) \end{aligned} \tag{49}$$

where the other minimum distances, except $d_{M,1}, d_{M,k}, d_{L,j}$, and $d_{L,l+1}$, are arbitrarily combined. We modify $f_{l+1,\sharp 1}$ such that $d_{M,1}$ and $d_{L,l+1}$ are combined, and $d_{M,k}$ and $d_{L,j}$ are combined. The modified f_{l+1} is denoted by $f_{l+1,\sharp 2}$:

$$\begin{aligned} f_{l+1,\sharp 2} = & \left(\frac{d_{M,1}^2}{2} + d_{M,1}d_{L,l+1} + d_{L,l+1}^2\right) + \left(\frac{d_{M,k}^2}{2} + d_{M,k}d_{L,j} + d_{L,j}^2\right) \\ & + \sum_{\substack{i=2 \\ i \neq k}}^{l+1} \left(\frac{\tilde{d}_{M,i}^2}{2} + \tilde{d}_{M,i}\tilde{d}_{L,i} + \tilde{d}_{L,i}^2\right). \end{aligned} \tag{50}$$

The difference between $f_{l+1,\sharp 1}$ and $f_{l+1,\sharp 2}$ is given by

$$f_{l+1,\sharp 1} - f_{l+1,\sharp 2} = (d_{M,1} - d_{M,k})(d_{L,j} - d_{L,l+1}) > 0 \tag{51}$$

because $d_{M,1} > d_{M,k}$ and $d_{L,j} > d_{L,l+1}$. From (51), $f_{l+1,\sharp 1}$, given by (49), cannot be the minimum of f_{l+1}, and thus the above assumption is false. We have thus showed that the largest distance for the MSBs, $d_{M,1}$ should be combined with the smallest distance for the LSBs, $d_{L,l+1}$. The other minimum distances, except $d_{M,1}$ and $d_{L,l+1}$, are given by

$$d_{M,2} > d_{M,3} > \cdots > d_{M,l+1} \quad \text{and} \quad d_{L,1} > d_{L,2} > \cdots > d_{L,l}. \tag{52}$$

By the induction hypothesis, the following is the minimum for $2l$ distances given by (52):

$$\sum_{i=1}^{l} \left(\frac{d_{M,i+1}^2}{2} + d_{M,i+1}d_{L,l+1-i} + d_{L,l+1-i}^2\right). \tag{53}$$

Thus, the minimum of f_{l+1} is given by

$$\begin{aligned} & \frac{d_{M,1}^2}{2} + d_{M,1}d_{L,l+1} + d_{L,l+1}^2 + \sum_{i=1}^{l} \left(\frac{d_{M,i+1}^2}{2} + d_{M,i+1}d_{L,l+1-i} + d_{L,l+1-i}^2\right) \\ & = \sum_{i=1}^{l+1} \left(\frac{d_{M,i}^2}{2} + d_{M,i}d_{L,l+2-i} + d_{L,l+2-i}^2\right). \end{aligned} \tag{54}$$

Setting $N = l+1$ in (46), we obtain $f^*_{l+1} = \sum_{i=1}^{l+1}(d^2_{M,i}/2 + d_{M,i}d_{L,l+2-i} + d^2_{L,l+2-i})$, and this is identical to (54). Hence, (46) holds for $N = l+1$.

□

Corollary 6 and Theorem 7 indicate that the average power of all the multiplexed constellations is minimized by permuting distances according to (45), while the BERs are unchanged for high SNR.

Next, we consider the peak signal power of the multiplexed hierarchical constellations. If we assume that all the hierarchical constellations are time-multiplexed, the peak power of all the multiplexed constellations, S_{peak}, is given by

$$S_{peak} = \max\left[\left\{S_{peak,i}\middle|1 \le i \le N\right\}\right] \tag{55}$$

where max$[X]$ denotes the maximum element of the set X, and $S_{peak,i}$ is the peak power of a hierarchical constellation i. For hierarchical 16 QAM, $S_{peak,i}$ is given by

$$S_{peak,i} = 2\left(\frac{d_{M,i}}{2} + d_{L,i}\right)^2 = \frac{d^2_{M,i}}{2} + 2d_{M,i}d_{L,i} + 2d^2_{L,i}. \tag{56}$$

Theorem 8: After the distances are permuted as described in Theorem 5, the peak power of all the multiplexed hierarchical 16 QAM constellations, S_{peak}, given by

$$S_{peak} = \max\left[\left\{S_{peak,i}\middle|1 \le i \le N\right\}\right] = \max\left[\left\{\frac{\tilde{d}^2_{M,i}}{2} + 2\tilde{d}_{M,i}\tilde{d}_{L,i} + 2\tilde{d}^2_{L,i}\middle|1 \le i \le N\right\}\right] \tag{57}$$

is minimized if the distances are permuted according to (45) of Theorem 7.

Proof: When (45) is satisfied, the corresponding S_{peak}, denoted by $S_{peak,\sharp 1}$, is given by

$$\begin{aligned} S_{peak,\sharp 1} &= \max\left[\left\{\frac{d^2_{M,i}}{2} + 2d_{M,i}d_{L,N+1-i} + 2d^2_{L,N+1-i}\middle|1 \le i \le N\right\}\right] \\ &= \frac{d^2_{M,j}}{2} + 2d_{M,j}d_{L,N+1-j} + 2d^2_{L,N+1-j}, \end{aligned} \tag{58}$$

for some j in the range of $1 \le j \le N$. We will contradict the following assumption: When distances are permuted in some way other than (45), the corresponding S_{peak}, denoted by $S_{peak,\sharp 2}$, is smaller than $S_{peak,\sharp 1}$. Let $d_{L,k}$ be the distance with which $d_{M,j}$ is combined (for some k in the range of $1 \le k \le N$) when the distances are permuted in a different manner from (45). The possible values of k can be classified into

$$1 \le k < N+1-j, \quad k = N+1-j, \quad \text{and} \quad N+1-j < k \le N. \tag{59}$$

i) For $1 \leq k < N+1-j$, $S_{peak,\sharp2} > S_{peak,\sharp1}$. To see this, note that

$$S_{peak,\sharp2} \geq \frac{1}{2}d_{M,j}^2 + 2d_{M,j}d_{L,k} + 2d_{L,k}^2 > \frac{1}{2}d_{M,j}^2 + 2d_{M,j}d_{L,N+1-j} + 2d_{L,N+1-j}^2 = S_{peak,\sharp1} \quad (60)$$

where the strict inequality follows from $d_{L,k} > d_{L,N+1-j}$ (since $k < N+1-j$).

ii) For $k = N+1-j$, $S_{peak,\sharp2} \geq S_{peak,\sharp1}$ since $S_{peak,\sharp2} \geq \frac{1}{2}d_{M,j}^2 + 2d_{M,j}d_{L,N+1-j} + 2d_{L,N+1-j}^2 = S_{peak,\sharp1}$.

iii) For $N+1-j < k \leq N$, $S_{peak,\sharp2} > S_{peak,\sharp1}$. This is proved as follows: Since $d_{M,j}$ is combined with $d_{L,k}$, other distances $\{d_{M,i}|1 \leq i \leq N,\ i \neq j\}$ should be combined with $\{d_{L,i}|1 \leq i \leq N,\ i \neq k\}$. Note that

$$\left|\left\{d_{M,i}\middle|1 \leq i < j\right\}\right| = j-1 \quad \text{and} \quad \left|\left\{d_{L,i}\middle|N+1-j < i \leq N,\ i \neq k\right\}\right| = j-2 \quad (61)$$

where $|X|$ denotes the cardinality of the set X, and the equality of the second expression follows from $N+1-j < k \leq N$. Since $j-1 > j-2$ in (61), at least one element of $\{d_{M,i}|1 \leq i < j\}$ should be combined with one element of $\{d_{L,i}|1 \leq i \leq N+1-j\}$. Suppose that $d_{M,p}$ is combined with $d_{L,q}$ for some $p \in \{1, \cdots, j-1\}$ and $q \in \{1, \cdots, N+1-j\}$. Then, we have

$$S_{peak,\sharp2} \geq \frac{1}{2}d_{M,p}^2 + 2d_{M,p}d_{L,q} + 2d_{L,q}^2 > \frac{1}{2}d_{M,j}^2 + 2d_{M,j}d_{L,N+1-j} + 2d_{L,N+1-j}^2 = S_{peak,\sharp1} \quad (62)$$

where the strict inequality follows from the fact that $d_{M,p} > d_{M,j}$ and $d_{L,q} \geq d_{L,N+1-j}$ (since $p < j$ and $q \leq N+1-j$). From i), ii), and iii), it is seen that there is no possible way of permuting distances which makes $S_{peak,\sharp2}$ smaller than $S_{peak,\sharp1}$. Therefore, the assumption below (58) is false.

□

Theorems 7 and 8 tell us that the permutation of the distances that minimizes the average power of all the multiplexed hierarchical constellations also, coincidentally, minimizes the peak power. Note that from (5) and (45), these optimally permuted distances satisfy

$$\tilde{d}_{M,1} > \cdots > \tilde{d}_{M,N} > \tilde{d}_{L,N} > \cdots > \tilde{d}_{L,1}. \quad (63)$$

Corollary 9: When the distances are optimally permuted according to (45) of Theorem 7, the BERs of the data classes satisfy $P_1 < P_2 < \cdots < P_{2N}$ for high SNR if class i and class $2N+1-i$ are mapped to the MSBs and LSBs of constellation i, respectively ($1 \leq i \leq N$).

Proof: The proof is similar to the proof of Corollary 2.

□

Fig. 2 (b) depicts the multilevel UEP system using multiplexed hierarchical 16 QAM constellations based on Corollary 9 for eight data classes ($N = 4$).

Next, we generalize to hierarchical $2^{2J}/2^{2K}$ $(K > J \geq 1)$ QAM constellations. Recall that $d_{M_n,i}$ denotes the minimum distance for the nth MSBs ($1 \leq n \leq K$) of a hierarchical 2^{2K} QAM constellation i. Hierarchical $2^{2J}/2^{2K}$ QAM has two distinct minimum Euclidian distances such that [31]

$$d_{M_n,i} = \begin{cases} d_{M_J,i}, & \text{for } 1 \leq n \leq J \\ d_{M_K,i}, & \text{for } J+1 \leq n \leq K. \end{cases} \tag{64}$$

The average power of a hierarchical $2^{2J}/2^{2K}$ QAM constellation i ($1 \leq i \leq N$) can be expressed, from (17) and (64), as the following:

$$S_{avg,i} = \sum_{u=1}^{J}\sum_{v=u}^{J} \mu_{uv} d_{M_J,i}^2 + \sum_{u=1}^{J}\sum_{v=J+1}^{K} \mu_{uv} d_{M_J,i} d_{M_K,i} + \sum_{u=J+1}^{K}\sum_{v=u}^{K} \mu_{uv} d_{M_K,i}^2. \tag{65}$$

Lemma 10: For high SNR, the BERs of a hierarchical $2^{2J}/2^{2K}$ QAM constellation i ($1 \leq i \leq N$) are given by

$$P_{M_n,i} \approx \begin{cases} \frac{1}{2^{K-n}} Q\left(\frac{d_{M_J,i}}{2}\sqrt{\frac{2\gamma_s}{S_{avg}}}\right), & \text{for } 1 \leq n \leq J \\ \frac{1}{2^{K-n}} Q\left(\frac{d_{M_K,i}}{2}\sqrt{\frac{2\gamma_s}{S_{avg}}}\right), & \text{for } J+1 \leq n \leq K \end{cases} \tag{66}$$

where S_{avg} is given by (1) and (65).

Proof: The BERs of a hierarchical 2^{2K} QAM constellation i, $P_{M_n,i}^{app}$ ($1 \leq n \leq K-1$), given by (18), can be rewritten as

$$\begin{aligned} P_{M_n,i}^{app} &= \frac{1}{2^{K-n}} Q\left(\left(\frac{d_{M_n,i}}{2} + \sum_{q=n+1}^{K}\left\lfloor \frac{2^{K-q}}{2^{K-q+1}}\right\rfloor d_{M_q,i}\right)\sqrt{\frac{2\gamma_s}{S_{avg}}}\right) \\ &\quad + \sum_{p=1}^{2^{K-n}-1} \frac{1}{2^{K-n}} Q\left(\left(\frac{d_{M_n,i}}{2} + \sum_{q=n+1}^{K}\left\lfloor \frac{p+2^{K-q}}{2^{K-q+1}}\right\rfloor d_{M_q,i}\right)\sqrt{\frac{2\gamma_s}{S_{avg}}}\right) \\ &= \frac{1}{2^{K-n}} Q\left(\frac{d_{M_n,i}}{2}\sqrt{\frac{2\gamma_s}{S_{avg}}}\right) \\ &\quad + \sum_{p=1}^{2^{K-n}-1} \frac{1}{2^{K-n}} Q\left(\left(\frac{d_{M_n,i}}{2} + \sum_{q=n+1}^{K}\left\lfloor \frac{p+2^{K-q}}{2^{K-q+1}}\right\rfloor d_{M_q,i}\right)\sqrt{\frac{2\gamma_s}{S_{avg}}}\right). \end{aligned} \tag{67}$$

From (67), we have

$$\sum_{q=n+1}^{K}\left\lfloor \frac{p+2^{K-q}}{2^{K-q+1}}\right\rfloor d_{M_q,i} \geq \sum_{q=n+1}^{K}\left\lfloor \frac{1+2^{K-q}}{2^{K-q+1}}\right\rfloor d_{M_q,i} \geq \left\lfloor \frac{1+2^{0}}{2^{1}}\right\rfloor d_{M_K,i} = d_{M_K,i}, \tag{68}$$

where the first inequality follows from $p \geq 1$ in (67). From (67) and (68), it is clear that the first Q-function term of (67) is the only term having the minimum distance of $d_{M_n,i}$ for the nth MSBs ($1 \leq n \leq K-1$). Also, for $P^{app}_{M_K,i}$ (i.e., $n = K$) given by (18), it is clear that the first Q-function term is the only term having the minimum distance of $d_{M_K,i}$. From the condition of approximation described in Lemma 3, it follows that the Q-function term having the minimum distance in $P^{app}_{M_K,i}$, given by (18), is the same as that in $P_{M_n,i}$, the exact BER. Therefore, from (64) and (67), (66) is derived.

□

From (66), the average BER for $n = 1, \cdots J$th MSBs of constellation i , denoted by $(P_{M_J,i})_{avg}$, is given by

$$(P_{M_J,i})_{avg} = \frac{1}{J}\sum_{n=1}^{J} P_{M_n,i} \approx A_J Q\left(\frac{d_{M_J,i}}{2}\sqrt{\frac{2\gamma_s}{S_{avg}}}\right) \tag{69}$$

where $A_J = \frac{1}{J}\sum_{n=1}^{J} 1/2^{K-n}$. Likewise, the average BER for $n = J+1, \cdots, K$th MSBs of constellation i, denoted by $(P_{M_K,i})_{avg}$, is given by

$$(P_{M_K,i})_{avg} = \frac{1}{K-J}\sum_{n=J+1}^{K} P_{M_n,i} \approx A_K Q\left(\frac{d_{M_K,i}}{2}\sqrt{\frac{2\gamma_s}{S_{avg}}}\right) \tag{70}$$

where $A_K = \frac{1}{K-J}\sum_{n=J+1}^{K} 1/2^{K-n}$. Similar to the average power given by (65), the peak power of a hierarchical $2^{2J}/2^{2K}$ QAM constellation i ($1 \leq i \leq N$) can be expressed as

$$\begin{aligned} S_{peak,i} &= \sum_{u=1}^{K}\sum_{v=u}^{K} \lambda_{uv} d_{M_u,i} d_{M_v,i} \\ &= \sum_{u=1}^{J}\sum_{v=u}^{J} \lambda_{uv} d^2_{M_J,i} + \sum_{u=1}^{J}\sum_{v=J+1}^{K} \lambda_{uv} d_{M_J,i} d_{M_K,i} + \sum_{u=J+1}^{K}\sum_{v=u}^{K} \lambda_{uv} d^2_{M_K,i} \end{aligned} \tag{71}$$

where the λ_{uv} are constants.

Theorem 11: Theorems 5, 7 and 8, and Corollaries 6 and 9 hold for hierarchical $2^{2J}/2^{2K}$ QAM when

i) $d_{M,i}$ and $d_{L,i}$ are replaced by $d_{M_J,i}$ and $d_{M_K,i}$, respectively, and $P_{M,i}$ and $P_{L,i}$ are replaced by $(P_{M_J,i})_{avg}$ and $(P_{M_K,i})_{avg}$, respectively.

ii) Eqs. (2) and (56) are replaced by (65) and (71), respectively.

Proof: From (69) and (70), $A_J < A_K$, since

$$A_J = \frac{1}{J}\sum_{n=1}^{J} \frac{1}{2^{K-n}} < \frac{1}{2^{K-J}} \quad \text{and} \quad A_K = \frac{1}{K-J}\sum_{n=J+1}^{K} \frac{1}{2^{K-n}} > \frac{1}{2^{K-J-1}}. \tag{72}$$

Hence, Theorem 5 and Corollary 6 hold for hierarchical $2^{2J}/2^{2K}$ QAM.

Since $\sum_{u=1}^{J}\sum_{v=u}^{J}\mu_{uv}$, $\sum_{u=1}^{J}\sum_{v=J+1}^{K}\mu_{uv}$ and $\sum_{u=J+1}^{K}\sum_{v=u}^{K}\mu_{uv}$ of (65) are coefficients, just as $1/2$, 1, and 1 of (2) are coefficients, Theorem 7 holds for hierarchical $2^{2J}/2^{2K}$ QAM. Likewise, $\sum_{u=1}^{J}\sum_{v=u}^{J}\lambda_{uv}$, $\sum_{u=1}^{J}\sum_{v=J+1}^{K}\lambda_{uv}$, and $\sum_{u=J+1}^{K}\sum_{v=u}^{K}\lambda_{uv}$ of (71) are coefficients, just as $1/2$, 2, and 2 of (56) are coefficients, and thus Theorem 8 holds for hierarchical $2^{2J}/2^{2K}$ QAM.

□

IV. Asymmetric Hierarchical QAM Constellation

While the proposed methods provide a large number of levels of UEP, the peak-to-average power ratio (PAPR) typically will be increased when hierarchical constellations having distinct minimum distances are time-multiplexed. To mitigate this effect, we design an asymmetric hierarchical QAM which reduces the PAPR without performance loss. From here onwards, we refer to conventional hierarchical QAM, which has been presented in Sections II and III, as symmetric hierarchical QAM, in order to distinguish it from asymmetric hierarchical QAM.

A. *Asymmetric Hierarchical* 2^{2K} $(K \geq 2)$ *QAM Constellation*

For an asymmetric hierarchical 2^{2K} QAM, the minimum distances for the inphase and quadrature components are different from each other. Similar to the previous sections, we first present asymmetric hierarchical 16 QAM, depicted in Fig. 4, as a simple example. The MSB i_1 for the inphase component determines the first cluster, and its minimum distance is $d_M^{A,I}$. The MSB q_1 for the quadrature component determines the second cluster within the first cluster that i_1 determined, and its minimum distance is $d_M^{A,Q}$. The LSB i_2 for the inphase component determines the third cluster, and its minimum distance is $d_L^{A,I}$, and the LSB q_2 for the quadrature component determines the specific signal point within the third cluster, and has minimum distance $d_L^{A,Q}$. Asymmetric hierarchical 16 QAM has three embedded subconstellations, and it provides four levels of UEP if $d_M^{A,I} > d_M^{A,Q} > d_L^{A,I} > d_L^{A,Q}$, which will be shown below in Corollary 13.

In order to provide $2N$ levels of UEP, we consider multiplexing $N/2$ (N is assumed to be even) asymmetric hierarchical 16 QAM constellations instead of N symmetric hierarchical 16 QAM constellations. The average power per symbol of all the multiplexed asymmetric constellations,

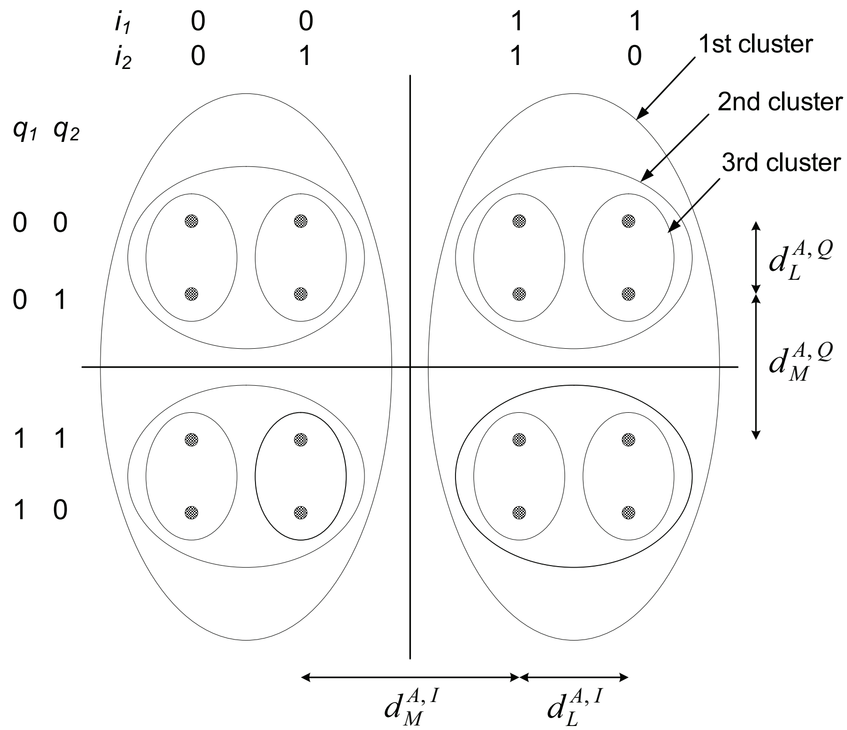

Fig. 4. Asymmetric hierarchical 16 QAM constellation.

S_{avg}^{A}, is given by

$$S_{avg}^{A} = \frac{1}{N/2} \sum_{i=1}^{N/2} S_{avg,i}^{A} \tag{73}$$

where $S_{avg,i}^{A}$ is the average power per symbol of asymmetric constellation i. For asymmetric hierarchical 16 QAM, $S_{avg,i}^{A}$ is given by

$$\begin{aligned} S_{avg,i}^{A} &= S_{avg,i}^{A,I} + S_{avg,i}^{A,Q} \\ &= \frac{1}{2}\left(\left(\frac{d_{M,i}^{A,I}}{2}\right)^2 + \left(\frac{d_{M,i}^{A,I}}{2} + d_{L,i}^{A,I}\right)^2\right) + \frac{1}{2}\left(\left(\frac{d_{M,i}^{A,Q}}{2}\right)^2 + \left(\frac{d_{M,i}^{A,Q}}{2} + d_{L,i}^{A,Q}\right)^2\right) \end{aligned} \tag{74}$$

where $S_{avg,i}^{A,I}$ and $S_{avg,i}^{A,Q}$ are the average powers per symbol for the inphase and quadrature components of asymmetric constellation i, respectively, and $d_{M,i}^{A,I}$, $d_{L,i}^{A,I}$, $d_{M,i}^{A,Q}$, and $d_{L,i}^{A,Q}$ are the minimum distances for the inphase MSB and LSB, and quadrature MSB and LSB, respectively. Note that the BERs of rectangular QAM are derived from those of the corresponding PAMs since the inphase and quadrature components are separated at the demodulator [31][33]. Let

$P_{M,i}^{A,I}$, $P_{L,i}^{A,I}$, $P_{M,i}^{A,Q}$, and $P_{L,i}^{A,Q}$ denote the BERs for the inphase MSB and LSB, and quadrature MSB and LSB of asymmetric hierarchical constellation i, respectively ($1 \leq i \leq N/2$). From (3), (73), and (74), they are derived as

$$\begin{aligned}
P_{M,i}^{A,I} &= \frac{1}{2}Q\left(\frac{d_{M,i}^{A,I}}{2}\sqrt{\frac{2\gamma_s}{S_{avg}^{A}}}\right)+\frac{1}{2}Q\left(\left(\frac{d_{M,i}^{A,I}}{2}+d_{L,i}^{A,I}\right)\sqrt{\frac{2\gamma_s}{S_{avg}^{A}}}\right)\\
P_{L,i}^{A,I} &= Q\left(\frac{d_{L,i}^{A,I}}{2}\sqrt{\frac{2\gamma_s}{S_{avg}^{A}}}\right)+\frac{1}{2}Q\left(\left(d_{M,i}^{A,I}+\frac{d_{L,i}^{A,I}}{2}\right)\sqrt{\frac{2\gamma_s}{S_{avg}^{A}}}\right)-\frac{1}{2}Q\left(\left(d_{M,i}^{A,I}+\frac{3d_{L,i}^{A,I}}{2}\right)\sqrt{\frac{2\gamma_s}{S_{avg}^{A}}}\right)\\
P_{M,i}^{A,Q} &= \frac{1}{2}Q\left(\frac{d_{M,i}^{A,Q}}{2}\sqrt{\frac{2\gamma_s}{S_{avg}^{A}}}\right)+\frac{1}{2}Q\left(\left(\frac{d_{M,i}^{A,Q}}{2}+d_{L,i}^{A,Q}\right)\sqrt{\frac{2\gamma_s}{S_{avg}^{A}}}\right)\\
P_{L,i}^{A,Q} &= Q\left(\frac{d_{L,i}^{A,Q}}{2}\sqrt{\frac{2\gamma_s}{S_{avg}^{A}}}\right)+\frac{1}{2}Q\left(\left(d_{M,i}^{A,Q}+\frac{d_{L,i}^{A,Q}}{2}\right)\sqrt{\frac{2\gamma_s}{S_{avg}^{A}}}\right)-\frac{1}{2}Q\left(\left(d_{M,i}^{A,Q}+\frac{3d_{L,i}^{A,Q}}{2}\right)\sqrt{\frac{2\gamma_s}{S_{avg}^{A}}}\right).
\end{aligned} \tag{75}$$

Theorem 12: Suppose there are N multiplexed symmetric hierarchical 16 QAM constellations whose minimum distances are given by $d_{M,1}, \cdots, d_{M,N}$ and $d_{L,1}, \cdots, d_{L,N}$. Also suppose there are $N/2$ asymmetric hierarchical 16 QAM constellations, and the minimum distances for the inphase and quadrature components of asymmetric hierarchical constellation i are the same as those of two distinct symmetric hierarchical constellations $x(i)$ and $y(i)$, respectively ($1 \leq i \leq N/2$). In other words,

$$d_{M,i}^{A,I} = d_{M,x(i)},\ d_{L,i}^{A,I} = d_{L,x(i)},\ d_{M,i}^{A,Q} = d_{M,y(i)},\ \text{and}\ d_{L,i}^{A,Q} = d_{L,y(i)} \quad (1 \leq i \leq N/2) \tag{76}$$

where $x(i)$ and $y(i)$ satisfy

$$x(i), y(i) \in \{1, \cdots, N\} \quad \text{and} \quad \{x(i), y(i) | 1 \leq i \leq N/2\} = \{1, \cdots, N\}. \tag{77}$$

With the minimum distances given by (76), the average power and BERs of $N/2$ multiplexed asymmetric hierarchical 16 QAM constellations are the same as those of N multiplexed symmetric hierarchical 16 QAM constellations, regardless of the choice of $x(i)$ and $y(i)$ satisfying (77).

Proof: From (74) and (76), $S_{avg,i}^{A}$ can be expressed as

$$\begin{aligned}
S_{avg,i}^{A} &= \frac{1}{2}\left(\left(\frac{d_{M,x(i)}}{2}\right)^2+\left(\frac{d_{M,x(i)}}{2}+d_{L,x(i)}\right)^2\right)+\frac{1}{2}\left(\left(\frac{d_{M,y(i)}}{2}\right)^2+\left(\frac{d_{M,y(i)}}{2}+d_{L,y(i)}\right)^2\right)\\
&= \frac{1}{2}S_{avg,x(i)}+\frac{1}{2}S_{avg,y(i)},
\end{aligned} \tag{78}$$

where the second equality follows from (2). From (73) and (78), S_{avg}^A is given by

$$S_{avg}^A = \frac{1}{N/2}\sum_{i=1}^{N/2}\left(\frac{1}{2}S_{avg,x(i)} + \frac{1}{2}S_{avg,y(i)}\right) = \frac{1}{N}\sum_{i=1}^{N/2}\Big(S_{avg,x(i)} + S_{avg,y(i)}\Big). \tag{79}$$

From (77), (79) can be rewritten as

$$S_{avg}^A = \frac{1}{N}\sum_{i=1}^{N} S_{avg,i} = S_{avg} \tag{80}$$

where the second equality follows from (1). We next compare the BERs of asymmetric and symmetric constellations. From (3), (75) and (76), we have

$$P_{M,i}^{A,I} = P_{M,x(i)},\ P_{L,i}^{A,I} = P_{L,x(i)},\ P_{M,i}^{A,Q} = P_{M,y(i)},\ \text{and}\ P_{L,i}^{A,Q} = P_{L,y(i)} \quad (1 \le i \le N/2). \tag{81}$$

From (77) and (81), a set of $2N$ BERs for $N/2$ multiplexed asymmetric constellations satisfy

$$\begin{aligned}\left\{P_{M,i}^{A,I}, P_{L,i}^{A,I}, P_{M,i}^{A,Q}, P_{L,i}^{A,Q}\middle| 1 \le i \le N/2\right\} &= \left\{P_{M,x(i)}, P_{L,x(i)}, P_{M,y(i)}, P_{L,y(i)}\middle| 1 \le i \le N/2\right\} \\ &= \left\{P_{M,i}, P_{L,i}\middle| 1 \le i \le N\right\}. \end{aligned} \tag{82}$$

Hence, a set of $2N$ BERs for $N/2$ multiplexed asymmetric constellations is the same as that for N multiplexed symmetric constellations.

□

Corollary 13: Suppose that the minimum distances of the N multiplexed symmetric hierarchical 16 QAM constellations satisfy (5) of Theorem 1. Then, with the minimum distances given by (76), $N/2$ multiplexed asymmetric hierarchical 16 QAM constellations also provide $2N$ levels of UEP.

Proof: Since $d_{M,i}$ and $d_{L,i}$ satisfy (5), $P_{M,i}$ and $P_{L,i}$ satisfy (4) by Theorem 1. From (82), it follows that $N/2$ multiplexed asymmetric hierarchical 16 QAM constellations also provide $2N$ levels of UEP.

□

As an example, suppose that there is single asymmetric hierarchical 16 QAM (i.e., $N = 2$), and $x(i)$ and $y(i)$ satisfying (77) are chosen as $x(1) = 1$ and $y(1) = 2$. From (76) and (81), (4) and (5) of Theorem 1 lead to the following:

$$P_{M,1}^{A,I} < P_{M,1}^{A,Q} < P_{L,1}^{A,I} < P_{L,1}^{A,Q} \quad \text{if} \quad d_{M,1}^{A,I} > d_{M,1}^{A,Q} > d_{L,1}^{A,I} > d_{L,1}^{A,Q}. \tag{83}$$

Next, we consider the peak power of all the multiplexed asymmetric hierarchical constellations, S_{peak}^{A}, which is given by

$$S_{peak}^{A} = \max\left[\left\{S_{peak,i}^{A}\middle|1 \leq i \leq N/2\right\}\right] \tag{84}$$

where $S_{peak,i}^{A}$ is the peak power of an asymmetric hierarchical constellation i. For asymmetric hierarchical 16 QAM, $S_{peak,i}^{A}$ is given by

$$S_{peak,i}^{A} = S_{peak,i}^{A,I} + S_{peak,i}^{A,Q} = \left(\frac{d_{M,i}^{A,I}}{2} + d_{L,i}^{A,I}\right)^2 + \left(\frac{d_{M,i}^{A,Q}}{2} + d_{L,i}^{A,Q}\right)^2 \tag{85}$$

where $S_{peak,i}^{A,I}$ and $S_{peak,i}^{A,Q}$ are the peak powers of the inphase and quadrature components of asymmetric hierarchical constellation i, respectively.

Theorem 14: Suppose that the minimum distances of the N multiplexed symmetric hierarchical 16 QAM satisfy (5) of Theorem 1. With the minimum distances given by (76), the peak power of all $N/2$ multiplexed asymmetric hierarchical 16 QAM constellations, S_{peak}^{A} , given by (84) and (85), is less than that of all N multiplexed symmetric hierarchical 16 QAM, S_{peak}, given by (55) and (56), regardless of the choice of $x(i)$ and $y(i)$ satisfying (77).

Proof: From (76) and (85), $S_{peak,i}^{A}$ is given by

$$S_{peak,i}^{A} = \left(\frac{d_{M,x(i)}}{2} + d_{L,x(i)}\right)^2 + \left(\frac{d_{M,y(i)}}{2} + d_{L,y(i)}\right)^2 = \frac{1}{2}S_{peak,x(i)} + \frac{1}{2}S_{peak,y(i)}, \tag{86}$$

where the second equality follows from (56). From (84) and (86), S_{peak}^{A} is given by

$$\begin{aligned} S_{peak}^{A} &= \max\left[\left\{\frac{1}{2}S_{peak,x(i)} + \frac{1}{2}S_{peak,y(i)}\middle|1 \leq i \leq N/2\right\}\right] \\ &= \frac{1}{2}S_{peak,x(j)} + \frac{1}{2}S_{peak,y(j)}, \end{aligned} \tag{87}$$

for some j in the range of $1 \leq j \leq N/2$. Since $x(i), y(i) \in \{1, \cdots, N\}$ from (77), we have

$$S_{peak,x(j)} \leq \max\left[\left\{S_{peak,i}\middle|1 \leq i \leq N\right\}\right] = S_{peak}, \text{ and } S_{peak,y(j)} \leq S_{peak}, \tag{88}$$

where the second equality of the first expression follows from (55). From (5) and (56), the peak powers of each symmetric hierarchical 16 QAM constellation satisfy

$$S_{peak,1} > S_{peak,2} > \cdots > S_{peak,N}. \tag{89}$$

From (77), (88) and (89), $S_{peak,x(j)}$ and $S_{peak,y(j)}$ satisfy either of the following:

$$S_{peak,x(j)} < S_{peak,y(j)} \leq S_{peak} \text{ or } S_{peak,y(j)} < S_{peak,x(j)} \leq S_{peak}. \tag{90}$$

From (87) and (90), we have

$$S_{peak}^{A} = \frac{1}{2} S_{peak,x(j)} + \frac{1}{2} S_{peak,y(j)} < S_{peak}. \tag{91}$$

□

Theorems 12 and 14 tell us that when asymmetric hierarchical 16 QAM is used instead of symmetric hierarchical 16 QAM, the PAPR is reduced without performance loss.

The following theorem states how to choose $x(i)$ and $y(i)$ $(1 \leq i \leq N/2)$ satisfying (77) to minimize the PAPR of all the multiplexed asymmetric hierarchical constellations.

Theorem 15: Suppose that the minimum distances of the N multiplexed symmetric hierarchical 16 QAM satisfy (5) of Theorem 1. Also suppose the minimum distances of $N/2$ multiplexed asymmetric hierarchical 16 QAM are given by (76). Then, from (84) and (86), S_{peak}^{A} is given by

$$S_{peak}^{A} = \max \left[\left\{ \frac{1}{2} S_{peak,x(i)} + \frac{1}{2} S_{peak,y(i)} \middle| 1 \leq i \leq N/2 \right\} \right] \tag{92}$$

and this is minimized if $x(i)$ and $y(i)$ satisfying (77) are chosen as

$$x(i) = i \quad \text{and} \quad y(i) = N + 1 - i \quad (1 \leq i \leq N/2). \tag{93}$$

Proof: The proof is similar to the proof of Theorem 8.

□

Next, we generalize to asymmetric hierarchical 2^{2K} $(K \geq 2)$ QAM. Let $d_{M_n,i}^{A,I}$ and $d_{M_n,i}^{A,Q}$ denote the minimum distances of the nth MSB $(1 \leq n \leq K)$ for the inphase and quadrature components of asymmetric hierarchical 2^{2K} QAM constellation i $(1 \leq i \leq N/2)$. From (17), the average power of asymmetric hierarchical 2^{2K} QAM constellation i, $S_{avg,i}^{A}$, can be expressed as

$$S_{avg,i}^{A} = S_{avg,i}^{A,I} + S_{avg,i}^{A,Q} = \sum_{u=1}^{K} \sum_{v=u}^{K} \frac{\mu_{uv}}{2} d_{M_u,i}^{A,I} d_{M_v,i}^{A,I} + \sum_{u=1}^{K} \sum_{v=u}^{K} \frac{\mu_{uv}}{2} d_{M_u,i}^{A,Q} d_{M_v,i}^{A,Q} \tag{94}$$

where $S_{avg,i}^{A,I}$ and $S_{avg,i}^{A,Q}$ are the average powers for the inphase and quadrature components of asymmetric constellation i.

Let $P_{M_n,i}^{A,I}$ and $P_{M_n,i}^{A,Q}$ denote the BERs of the nth MSB $(1 \leq n \leq K)$ for the inphase and quadrature components of asymmetric hierarchical 2^{2K} QAM constellation i $(1 \leq i \leq N/2)$. Recall that $P_{M_n,i}$ denotes the BER of the nth MSBs $(1 \leq n \leq K)$ of symmetric hierarchical 2^{2K} QAM constellation i $(1 \leq i \leq N)$.

Theorem 16: Suppose that there are N multiplexed symmetric hierarchical 2^{2K} QAM whose minimum distances are given by $d_{M_n,1}, \cdots, d_{M_n,N}$ $(1 \leq n \leq K)$. Also suppose that the minimum distances of $N/2$ multiplexed asymmetric hierarchical 2^{2K} QAM satisfy

$$d^{A,I}_{M_n,i} = d_{M_n,x(i)} \quad \text{and} \quad d^{A,Q}_{M_n,i} = d_{M_n,y(i)} \quad (1 \leq n \leq K,\ 1 \leq i \leq N/2) \tag{95}$$

where $x(i)$ and $y(i)$ satisfy (77). Theorem 12 holds for asymmetric hierarchical 2^{2K} QAM when

i) $d^{A,I}_{M,i}$ and $d^{A,I}_{L,i}$ are replaced by $d^{A,I}_{M_n,i}$ $(1 \leq n \leq K)$; $d^{A,Q}_{M,i}$ and $d^{A,Q}_{L,i}$ are replaced by $d^{A,Q}_{M_n,i}$ $(1 \leq n \leq K)$; $d_{M,i}$ and $d_{L,i}$ are replaced by $d_{M_n,i}$ $(1 \leq n \leq K)$.

ii) $P^{A,I}_{M,i}$ and $P^{A,I}_{L,i}$ are replaced by $P^{A,I}_{M_n,i}$ $(1 \leq n \leq K)$; $P^{A,Q}_{M,i}$ and $P^{A,Q}_{L,i}$ are replaced by $P^{A,Q}_{M_n,i}$ $(1 \leq n \leq K)$; $P_{M,i}$ and $P_{L,i}$ are replaced by $P_{M_n,i}$ $(1 \leq n \leq K)$.

iii) Eq. (76) is replaced by (95).

Proof: We omit the proof for conciseness, but it can be found in [32].

□

We next consider the peak power for asymmetric hierarchical 2^{2K} QAM. In the following, we rewrite the peak power of symmetric hierarchical 2^{2K} QAM constellation i $(1 \leq i \leq N)$, $S_{peak,i}$, given by (71):

$$S_{peak,i} = \sum_{u=1}^{K} \sum_{v=u}^{K} \lambda_{uv} d_{M_u,i} d_{M_v,i}. \tag{96}$$

From (96), the peak power of asymmetric hierarchical 2^{2K} QAM constellation i, $S^{A}_{peak,i}$, can be expressed as

$$S^{A}_{peak,i} = S^{A,I}_{peak,i} + S^{A,Q}_{peak,i} = \sum_{u=1}^{K} \sum_{v=u}^{K} \frac{\lambda_{uv}}{2} d^{A,I}_{M_u,i} d^{A,I}_{M_v,i} + \sum_{u=1}^{K} \sum_{v=u}^{K} \frac{\lambda_{uv}}{2} d^{A,Q}_{M_u,i} d^{A,Q}_{M_v,i} \tag{97}$$

where $S^{A,I}_{peak,i}$ and $S^{A,Q}_{peak,i}$ are the peak powers for the inphase and quadrature components of asymmetric constellation i.

Theorem 17: Theorems 14 and 15 hold for asymmetric hierarchical 2^{2K} QAM when

i) $d^{A,I}_{M,i}$ and $d^{A,I}_{L,i}$ are replaced by $d^{A,I}_{M_n,i}$ $(1 \leq n \leq K)$; $d^{A,Q}_{M,i}$ and $d^{A,Q}_{L,i}$ are replaced by $d^{A,Q}_{M_n,i}$ $(1 \leq n \leq K)$; $d_{M,i}$ and $d_{L,i}$ are replaced by $d_{M_n,i}$ $(1 \leq n \leq K)$.

ii) $S_{peak,i}$ given by (56) is replaced by (96).

iii) $S^{A}_{peak,i}$ given by (85) is replaced by (97).

iv) Eq. (5) of Theorem 1 is replaced by (20) of Theorem 4.

v) Eq. (76) is replaced by (95).

Proof: We omit the proof for conciseness, but it can be found in [32].

□

We note that, like other rectangular QAM constellations, the asymmetric hierarchical 2^{2K} QAM can be easily generated as two PAM signals impressed on the inphase and quadrature carriers, and possesses the distinct advantage of being easily demodulated. Hence, it does not increase any decoding complexities, compared to conventional hierarchical or non-hierarchical rectangular QAM constellations.

V. Multilevel UEP Based on Multiplexing Hierarchical QAM Constellations Having Constant Power

In this section, we consider the case where it is desirable for the multiplexed hierarchical QAM constellations to have the same average power (i.e., constant power), either due to the limited capability of a power amplifier, or for cochannel interference control.

A. Symmetric Hierarchical $2^{2J}/2^{2K}$ $(K > J \geq 1)$ QAM Constellation

Theorem 18: When N multiplexed symmetric hierarchical 16 QAM constellations are required to have constant power, there exist minimum distances satisfying

$$d_{M,1} > d_{M,2} > \cdots > d_{M,N} > d_{L,N} > d_{L,N-1} > \cdots > d_{L,1}. \tag{98}$$

Proof: The proof of this theorem as well as the proofs of all other theorems in this section are not included here for conciseness, but they can be found in [32].

□

From (63) and (98), it is seen that even if symmetric hierarchical 16 QAM constellations have constant power, the suggested UEP system, depicted in Fig. 2 (b), can provide $2N$ levels of UEP for high SNR.

Theorem 18 holds for symmetric hierarchical $2^{2J}/2^{2K}$ $(K > J \geq 1)$ QAM, when $d_{M,i}$ and $d_{L,i}$ are replaced by $d_{M_J,i}$ and $d_{M_K,i}$, respectively.

B. Asymmetric Hierarchical 2^{2K} $(K \geq 2)$ QAM Constellation

Theorem 19: Suppose that $N/2$ multiplexed asymmetric hierarchical 16 QAM constellations are required to have constant power, and their minimum distances are given by (76). If $x(i)$ and

$y(i)$ are chosen according to (93) of Theorem 15, there exist minimum distances satisfying both (5) of Theorem 1 and (76).

□

From Corollary 13 and Theorem 19, it follows that even if asymmetric hierarchical 16 QAM constellations have constant power, $2N$ levels of UEP can be achieved.

Theorem 19 holds for asymmetric hierarchical 2^{2K} $(K \geq 3)$ QAM, when

i) $d_{M,i}$ and $d_{L,i}$ are replaced by $d_{M_n,i}$ ($1 \leq n \leq K$).

ii) Eq. (76) is replaced by (95).

iii) Eq. (5) of Theorem 1 is replaced by (20) of Theorem 4.

Theorem 20: Suppose that $N/2$ multiplexed asymmetric hierarchical 2^{2K} $(K \geq 2)$ QAM constellations are required to have constant power. Then the performance of the system stays the same or degrades compared to the case where multiplexed constellations are not required to have constant power.

□

VI. The Performance of the Proposed UEP System for Progressive Bitstream Transmission

In this section, we analyze the performance of the proposed UEP system for progressive image source transmission over Rayleigh fading channels. We first consider the UEP system depicted in Fig. 2 (a). The system takes successive blocks (data classes) of the compressed progressive bitstream, and transforms them into a sequence of channel codewords of fixed length l_c [22] with error detection and correction capability. Then, the coded classes are mapped to the multiplexed symmetric hierarchical 16 QAM constellations. At the receiver, if a received class is correctly decoded, then the next class is considered by the decoder. Otherwise, the decoding is stopped and the image is reconstructed from the correctly decoded classes. We assume that all decoding errors can be detected.

Let r_i be an error correction code rate for class i ($1 \leq i \leq 2N$), and $\underline{d}_i = (d_{M,c(i)},\ d_{L,c(i)})$ be a pair of minimum distances of some specific constellation $c(i)$ ($1 \leq c(i) \leq N$) to which class i ($1 \leq i \leq 2N$) is mapped. From Corollary 2, $\underline{d}_i$ ($1 \leq i \leq 2N$) is given by

$$\underline{d}_i = \begin{cases} (d_{M,i},\ d_{L,i}), & \text{for } 1 \leq i \leq N \\ (d_{M,i-N},\ d_{L,i-N}), & \text{for } N+1 \leq i \leq 2N \end{cases} \tag{99}$$

where $d_{M,1}, \cdots, d_{M,N}$ and $d_{L,1}, \cdots, d_{L,N}$ satisfy (5) of Theorem 1 to achieve $2N$ levels of UEP. Let $p(r_i, \underline{d}_i, \gamma_s)$ denote the probability of a decoding error of class i. Then, the probability that no decoding errors occur in the first i classes with an error in the next one, $P_{c,i}$ is given by

$$P_{c,i} = p(r_{i+1}, \underline{d}_{i+1}, \gamma_s) \prod_{j=1}^{i} \left(1 - p(r_j, \underline{d}_j, \gamma_s)\right) \quad \text{for } 1 \le i \le 2N-1. \tag{100}$$

Note that $P_{c,0} = p(r_1, \underline{d}_1, \gamma_s)$ is the probability of an error in the first class, and $P_{c,2N} = \prod_{j=1}^{2N} \left(1 - p(r_j, \underline{d}_j, \gamma_s)\right)$ is the probability that all $2N$ classes are correctly decoded. The end-to-end performance can be measured by the expected distortion, $E[D]$, given by

$$E[D] = \sum_{i=0}^{2N} P_{c,i} D_i \tag{101}$$

where D_i is the reconstruction error using the first i classes ($1 \le i \le 2N$), and D_0 is a constant. For the case of an uncoded system, D_i is given by $D_i = V(il_c)$, where $V(x)$ denotes the operational rate-distortion function of the source coder. Also, for the uncoded system, the probability of a decoding error of class i, $p(r_i, \underline{d}_i, \gamma_s) = p(\underline{d}_i, \gamma_s)$, can be obtained analytically:

$$p(\underline{d}_i, \gamma_s) = 1 - \left\{1 - P_i(\underline{d}_i, \gamma_s)\right\}^{l_c}. \tag{102}$$

Recall that P_i, a function of $\underline{d}_i$ and γ_s, is the BER of data class i. P_i ($1 \le i \le 2N$) is given by (3) and (15) of Corollary 2. We define a frame as a group of constellation symbols to which one image bitstream is mapped. We assume the channel experiences slow Rayleigh fading such that the fading coefficients are nearly constant over a frame. With this channel model, from (100)–(102), the expected distortion for the uncoded system is given by

$$\begin{aligned} E[D] = \int_0^\infty \Bigg\{ & \left(1 - \left\{1 - P_1(\underline{d}_1, h^2\gamma_s)\right\}^{l_c}\right) V(0) \\ & + \sum_{i=1}^{2N-1} \left[\left(1 - \left\{1 - P_{i+1}(\underline{d}_{i+1}, h^2\gamma_s)\right\}^{l_c}\right) \prod_{j=1}^{i} \left\{1 - P_j(\underline{d}_j, h^2\gamma_s)\right\}^{l_c} \right] V(il_c) \\ & + \prod_{j=1}^{2N} \left\{1 - P_j(\underline{d}_j, h^2\gamma_s)\right\}^{l_c} V(2Nl_c) \Bigg\} f(h) dh \end{aligned} \tag{103}$$

where h is the Rayleigh-distributed envelope of complex channel coefficients and $f(h)$ is the Rayleigh-distributed probability density function of h. Note that for a given SNR of γ_s, $E[D]$ is the conditional expected distortion. In situations when exact SNR information is not available at the transmitter, one can find the minimum distances, $\underline{d}_1, \cdots, \underline{d}_{2N}$ (or $d_{M,1}, \cdots, d_{M,N}$ and

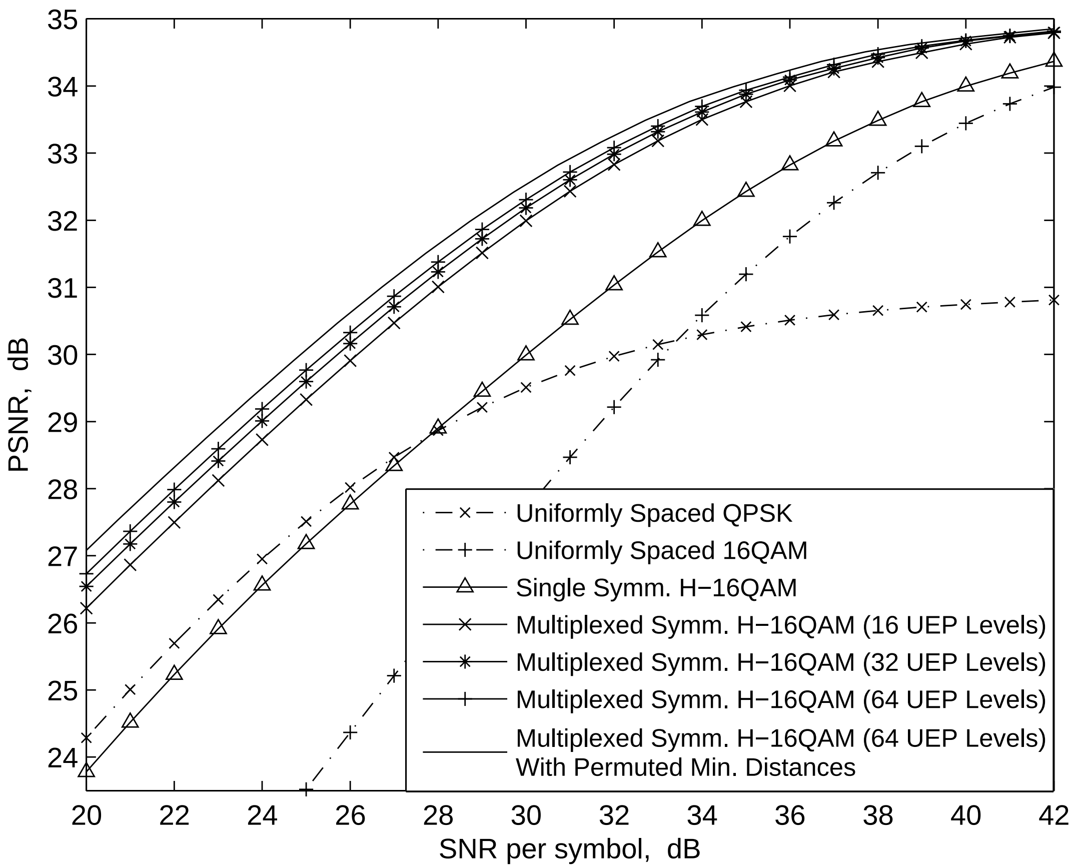

Fig. 5. PSNR performance of UEP system using multiplexed symmetric hierarchical 16 QAM (H-16QAM denotes hierarchical 16 QAM).

$d_{L,1}, \cdots, d_{L,N}$), which minimize the expected distortion over a range of expected SNRs using the weighted cost function

$$\underset{\underline{d}_1,\cdots,\underline{d}_{2N}}{\arg\min} \frac{\int_0^\infty \omega(\gamma_s)E[D]d\gamma_s}{\int_0^\infty \omega(\gamma_s)d\gamma_s} \tag{104}$$

where $\omega(\gamma_s)$ in [0, 1] is the weight function. For example, $\omega(\gamma_s)$ can be given by

$$\omega(\gamma_s) = \begin{cases} 1, & \text{for } \gamma_s^a \le \gamma_s \le \gamma_s^b \\ 0, & \text{otherwise.} \end{cases} \tag{105}$$

TABLE I

PAPR OF MULTIPLEXED SYMMETRIC OR ASYMMETRIC HIERARCHICAL 16 QAM

	PAPR (dB)		
Number of UEP levels	4	16	64
Multiplexed symmetric hierarchical 16 QAM	3.31	6.87	9.43
Multiplexed symmetric hierarchical 16 QAM with permuted min. distances	2.82	5.84	8.32
Multiplexed asymmetric hierarchical 16 QAM	1.11	4.18	6.60
Multiplexed asymmetric hierarchical 16 QAM having constant power	1.11	1.43	1.46

TABLE II

PAPR OF UNIFORMLY SPACED 16 QAM AND SINGLE SYMMETRIC HIERARCHICAL 16 QAM

	PAPR (dB)
Uniformly spaced 16 QAM	2.55
Single symmetric hierarchical 16 QAM	0.90

VII. NUMERICAL RESULTS

We evaluate the performance of the proposed UEP system using multiplexed hierarchical 16 QAM constellations for the progressive source coder SPIHT [26] as an example. We provide the results for the standard 8 bits per pixel (bpp) 512×512 Lena image with a transmission rate of 0.375 bpp. To compare the image quality, we use peak-signal-to-noise ratio (PSNR) defined as

$$\text{PSNR} = 10\log\frac{255^2}{E[D]} \quad \text{(dB)} \tag{106}$$

where 255 is due to the 8 bpp image, and $E[D]$ is given by (103).

We present the PSNR performance for the uncoded case by numerically evaluating (103)–(106) as follows: We first compute (104) for the block Rayleigh fading channel using the expected distortion, $E[D]$, given by (103), and the weight function, $\omega(\gamma_s)$, given by (105). Next, with

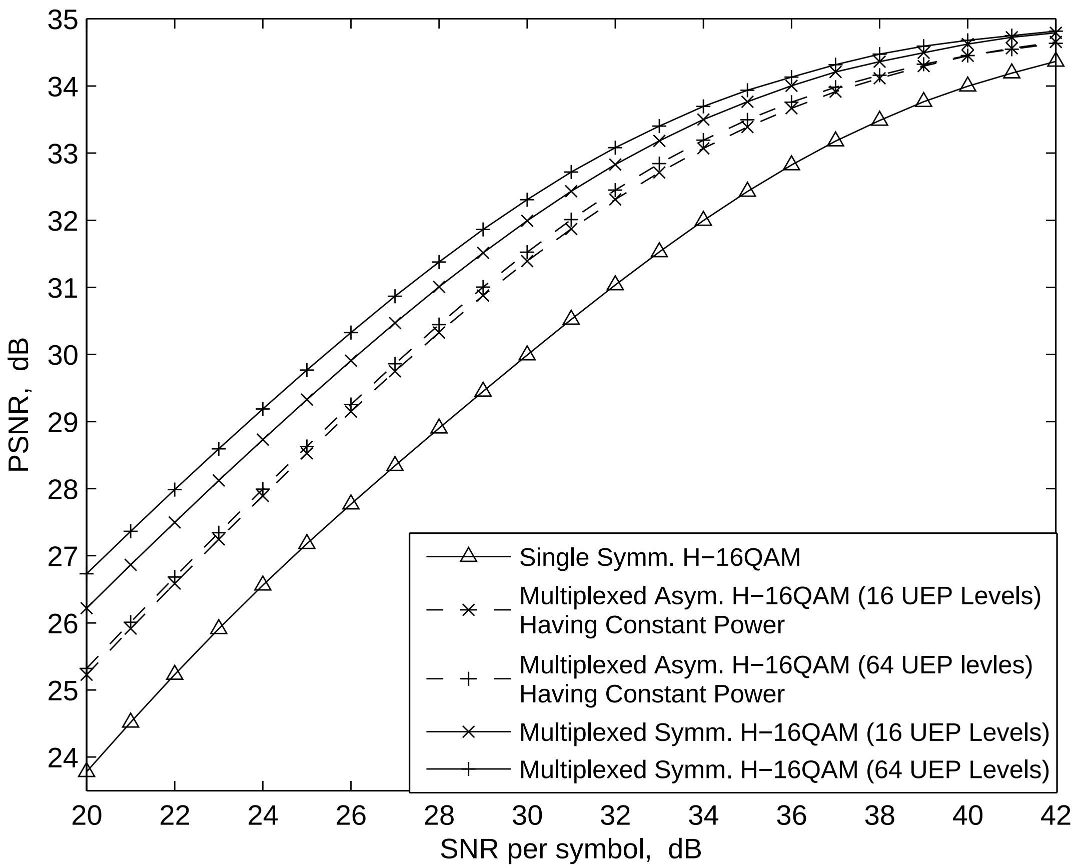

Fig. 6. PSNR performance of UEP system using multiplexed asymmetric hierarchical 16 QAM having constant power (H-16QAM denotes hierarchical 16 QAM).

$\underline{d}_1, \cdots, \underline{d}_{2N}$ (or $d_{M,1}, \cdots, d_{M,N}$ and $d_{L,1}, \cdots, d_{L,N}$) obtained from (104), we evaluate PSNR using (103) and (106) over a range of expected SNRs given by (105).

Fig. 5 shows the PSNR performance of the multiplexed symmetric hierarchical 16 QAM constellations. For reference, it also shows PSNRs for single symmetric hierarchical 16 QAM, as well as uniformly spaced QPSK and 16 QAM constellations. The PSNR of single symmetric hierarchical constellation is evaluated in the same way as that for multiplexed symmetric hierarchical constellations. From Fig. 5, it is seen that multiplexed symmetric hierarchical constellations improve the performance more than does single symmetric hierarchical constellation. It is also seen that 32 multiplexed symmetric hierarchical 16 QAM constellations, which provide 64 levels of UEP, have almost saturated performance in this evaluation. However, by optimally permuting the minimum distances according to Theorem 7, an additional SNR gain of more than 0.5 dB is achieved. Note that the performance of $N/2$ multiplexed asymmetric hierarchical constellations is the same as that of N multiplexed symmetric hierarchical constellations (N=8,16,32) as stated by Theorem 12, though the former is not depicted here.

Table I shows the PAPRs of the multiplexed symmetric or asymmetric hierarchical 16 QAM constellations. For reference, the PAPRs of single symmetric hierarchical 16 QAM and uniformly spaced 16 QAM constellations are also listed in Table II. From Tables I and II, it is seen that when symmetric hierarchical 16 QAM constellations are time-multiplexed, they have larger PAPR than does uniformly spaced 16 QAM as well as single symmetric hierarchical 16 QAM constellation. Table I also shows that PAPR is reduced when asymmetric hierarchical constellation is used, as stated by Theorem 14.

Fig. 6 shows the PSNR performance of the multiplexed asymmetric hierarchical 16 QAM constellations having constant power. It is shown that the performance is degraded when constellations are required to have constant power, which is consistent with Theorem 20. However, as seen from Table I, this scheme provides PAPR smaller than uniformly-spaced QAM, and a high PAPR problem is solved.

VIII. CONCLUSION

Progressive image or scalable video encoders employ progressive transmission, so that encoded data have gradual differences of importance in their bitstreams, which necessitates multiple levels of UEP. Though hierarchical modulation has been intensively studied for digital broadcasting or multimedia transmission, methods of achieving a large number of levels of UEP for progressive mode of transmission have rarely been studied.

In this paper, we proposed a multilevel UEP system using multiplexed hierarchical modulation for progressive transmission over mobile radio channels. Specifically, we proposed a way of multiplexing N hierarchical 2^{2K} QAM constellations ($K \geq 2$) and proved that KN levels of UEP are achieved, under the assumption that the SNR of interest for the nth most important bits is reasonably large so that the probability of noise exceeding the Euclidian distance of $d_{M_{n-1}} + \frac{1}{2}d_{M_n}$ is insignificant compared to that of noise exceeding $\frac{1}{2}d_{M_n}$, where d_{M_n} and $d_{M_{n-1}}$ are the minimum distances for the nth and $n-1$th important bits, respectively ($2 \leq n \leq K$).

This assumption is based on the fact that for hierarchical constellations, the minimum distance for more important bits is greater than that for less important bits (i.e., $d_{M_{n-1}} > d_{M_n}$). As a special case, for hierarchical 16 QAM ($K = 2$), we showed that $2N$ levels of UEP are achieved without the assumption.

When the BER is dominated by the Q-function term having the minimum Euclidian distance, we derived an optimal multiplexing approach which minimizes both the average and peak powers for hierarchical $2^{2J}/2^{2K}$ QAM ($K > J \geq 1$) constellations (typical examples are 4/16 QAM and 4/64 QAM which are employed in the DVB-T standard). While the suggested methods achieve multiple levels of UEP, the PAPR typically will be increased when constellations having distinct minimum distances are time-multiplexed. To mitigate this effect, an asymmetric hierarchical QAM constellation, which reduces the PAPR without performance loss, was proposed. We also considered the case where multiplexed constellations need to have constant power, and showed that multilevel UEP can be achieved while the performance stays the same or degrades in this case. Numerical results showed that the proposed multilevel UEP system based on multiplexed modulation significantly enhances the performance for progressive transmission over Rayleigh fading channels without any additional system bandwidth or transmit power.

APPENDIX A
PROOF OF LEMMA 3

A. *Gray coded bit mapping vector for hierarchical 2^K PAM*

For a hierarchical 2^K PAM constellation, let $g_{n,i}$ denote the Gray code for the nth MSB ($1 \leq n \leq K$) assigned to the ith signal point ($1 \leq i \leq 2^K$) from the left. Then, it can be shown that the 2^K-tuple Gray coded bit mapping vector, $\mathbf{g}_n = \left[g_{n,1}\ g_{n,2} \cdots g_{n,2^K}\right]$, for the nth MSB is given by

$$\mathbf{g}_n = \begin{cases} \left[\mathbf{0}_{2^{K-1}}\ \mathbf{1}_{2^{K-1}}\right], & \text{for } n = 1 \\ \left[\mathbf{0}_{2^{K-n}}\ \mathbf{1}_{2^{K-n}}\ \mathbf{1}_{2^{K-n}}\ \mathbf{0}_{2^{K-n}} \cdots \mathbf{0}_{2^{K-n}}\ \mathbf{1}_{2^{K-n}}\ \mathbf{1}_{2^{K-n}}\ \mathbf{0}_{2^{K-n}}\right], & \text{for } 2 \leq n \leq K \end{cases} \tag{107}$$

where $\mathbf{0}_l$ is a l-tuple all zero vector, and $\mathbf{1}_l$ is a l-tuple all one vector.

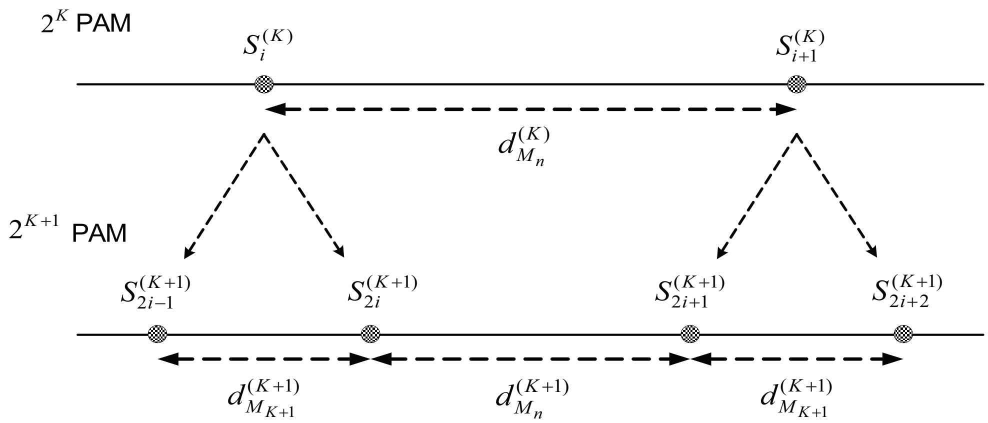

Fig. 7. The construction of hierarchical 2^{K+1} PAM from hierarchical 2^K PAM.

B. Euclidian distance between adjacent signal points for hierarchical 2^K PAM

Let $S_i^{(K)}$ $(1 \leq i \leq 2^K)$ and $S_i^{(K+1)}$ $(1 \leq i \leq 2^{K+1})$ denote the ith signal point from the left for hierarchical 2^K and 2^{K+1} PAM constellations, respectively. Also, let $d_{M_n}^{(K)}$ $(1 \leq n \leq K)$ and $d_{M_n}^{(K+1)}$ $(1 \leq n \leq K+1)$ denote minimum distances for the nth MSB of hierarchical 2^K and 2^{K+1} PAM constellations, respectively. Fig. 7 shows how hierarchical 2^{K+1} PAM is constructed from hierarchical 2^K PAM. There are two rules with regard to the construction of hierarchical 2^{K+1} PAM from hierarchical 2^K PAM:

i) The ith signal point for 2^K PAM, $S_i^{(K)}$, is replaced by the $2i-1$th and $2i$th signal points for 2^{K+1} PAM, $S_{2i-1}^{(K+1)}$ and $S_{2i}^{(K+1)}$, which satisfy

$$d\big(S_{2i-1}^{(K+1)}, S_{2i}^{(K+1)}\big) = d_{M_{K+1}}^{(K+1)} \quad \text{for } 1 \leq i \leq 2^K \tag{108}$$

where $d(X,Y)$ is the Euclidian distance between two signal points, X and Y.

ii) If the distance between $S_i^{(K)}$ and $S_{i+1}^{(K)}$ for 2^K PAM is $d_{M_n}^{(K)}$, then the distance between $S_{2i}^{(K+1)}$ and $S_{2i+1}^{(K+1)}$ for 2^{K+1} PAM is $d_{M_n}^{(K+1)}$. That is, for $1 \leq i \leq 2^K - 1$ and $1 \leq n \leq K$,

$$d\big(S_{2i}^{(K+1)}, S_{2i+1}^{(K+1)}\big) = d_{M_n}^{(K+1)} \quad \text{if} \quad d\big(S_i^{(K)}, S_{i+1}^{(K)}\big) = d_{M_n}^{(K)}. \tag{109}$$

As an example, Fig. 8 depicts hierarchical 4 and 8 PAM constellations.

We will prove the following by induction: For hierarchical 2^K PAM ($K \geq 2$), the Euclidian distance between adjacent signal points is given by

$$d\big(S_{(2i-1)2^{K-n}}^{(K)}, S_{(2i-1)2^{K-n}+1}^{(K)}\big) = d_{M_n}^{(K)} \quad \text{for } 1 \leq i \leq 2^{n-1} \text{ and } 1 \leq n \leq K. \tag{110}$$

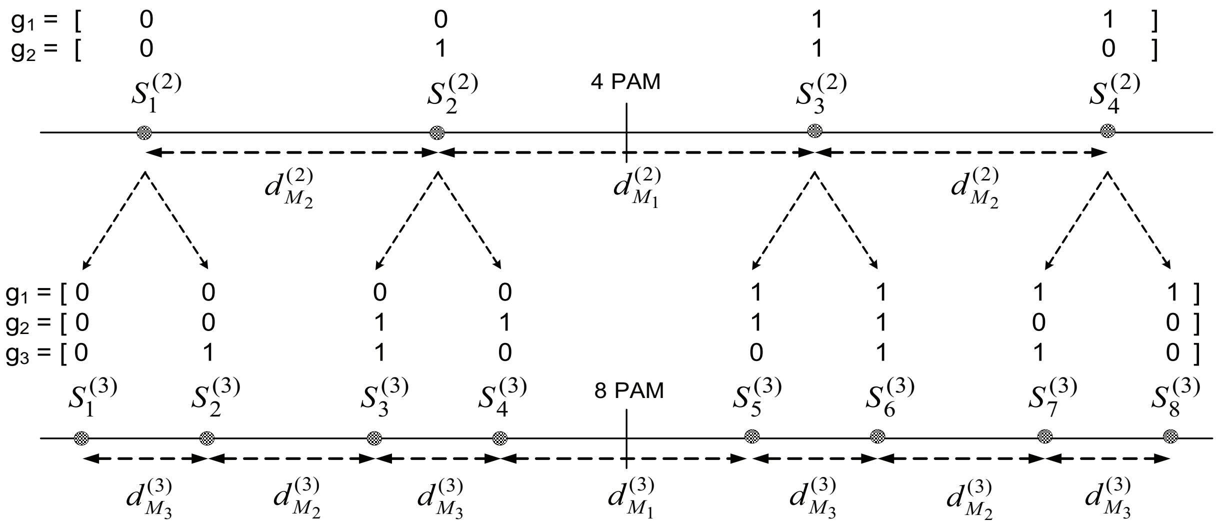

Fig. 8. Hierarchical 4 and 8 PAM constellations.

Consider hierarchical 4 PAM. From Fig. 8, it is seen that

$$d\big(S_2^{(2)}, S_3^{(2)}\big) = d_{M_1}^{(2)} \quad \text{and} \quad d\big(S_1^{(2)}, S_2^{(2)}\big) = d\big(S_3^{(2)}, S_4^{(2)}\big) = d_{M_2}^{(2)}. \tag{111}$$

If we let $K = 2$ in (110), we have

$$d\big(S_{(2i-1)2^{2-n}}^{(2)}, S_{(2i-1)2^{2-n}+1}^{(2)}\big) = d_{M_n}^{(2)} \quad \text{for } 1 \le i \le 2^{n-1} \text{ and } 1 \le n \le 2. \tag{112}$$

From (112), for $n = 1$, we have

$$d\big(S_{(2i-1)2}^{(2)}, S_{(2i-1)2+1}^{(2)}\big) = d_{M_1}^{(2)} \quad \text{for } i = 1 \;\Leftrightarrow\; d\big(S_2^{(2)}, S_3^{(2)}\big) = d_{M_1}^{(2)} \tag{113}$$

where $A \Leftrightarrow B$ denotes A and B are identical. From (112), for $n = 2$, we have

$$d\big(S_{2i-1}^{(2)}, S_{2i}^{(2)}\big) = d_{M_2}^{(2)} \quad \text{for } i = 1, 2 \;\Leftrightarrow\; d\big(S_1^{(2)}, S_2^{(2)}\big) = d\big(S_3^{(2)}, S_4^{(2)}\big) = d_{M_2}^{(2)}. \tag{114}$$

It is seen that (113) and (114) are identical to (111). Suppose that (110) holds for 2^l PAM. That is,

$$d\big(S_{(2i-1)2^{l-n}}^{(l)}, S_{(2i-1)2^{l-n}+1}^{(l)}\big) = d_{M_n}^{(l)} \quad \text{for } 1 \le i \le 2^{n-1} \text{ and } 1 \le n \le l. \tag{115}$$

Consider hierarchical 2^{l+1} PAM. Eq. (109) can be rewritten as

$$d\big(S_{2i}^{(l+1)}, S_{2i+1}^{(l+1)}\big) = d_{M_n}^{(l+1)} \quad \text{if} \quad d\big(S_i^{(l)}, S_{i+1}^{(l)}\big) = d_{M_n}^{(l)}, \tag{116}$$

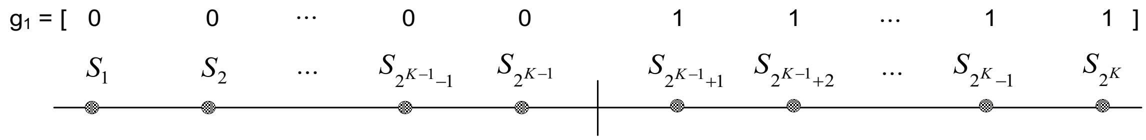

Fig. 9. Hierarchical 2^K PAM constellation with the bit mapping vector $\mathbf{g}_1$ for the MSB.

for $1 \leq i \leq 2^l - 1$ and $1 \leq n \leq l$. From (115) and (116), it can be shown that

$$d\big(S^{(l+1)}_{(2i-1)2^{l+1-n}}, S^{(l+1)}_{(2i-1)2^{l+1-n}+1}\big) = d^{(l+1)}_{M_n} \quad \text{for } 1 \leq i \leq 2^{n-1} \text{ and } 1 \leq n \leq l. \tag{117}$$

Eq. (108) can be rewritten as

$$d\big(S^{(l+1)}_{2i-1}, S^{(l+1)}_{2i}\big) = d^{(l+1)}_{M_{l+1}} \quad \text{for } 1 \leq i \leq 2^l. \tag{118}$$

From (118), (117) can be extended to the case $n = l + 1$. That is,

$$d\big(S^{(l+1)}_{(2i-1)2^{l+1-n}}, S^{(l+1)}_{(2i-1)2^{l+1-n}+1}\big) = d^{(l+1)}_{M_n} \quad \text{for } 1 \leq i \leq 2^{n-1} \text{ and } 1 \leq n \leq l+1. \tag{119}$$

If we let $K = l + 1$ in (110), it is identical to (119). Hence, (110) holds for hierarchical 2^{l+1} PAM.

For convenience, from here onwards, we use S_i and d_{M_n} instead of $S_i^{(K)}$ and $d_{M_n}^{(K)}$ for hierarchical 2^K PAM. For integers j, n in the range of $1 \leq j \leq 2^K - 1$ and $1 \leq n \leq K$, we define a function $f_n(j)$ as

$$f_n(j) = \begin{cases} 1, & \text{for } \ j = (2 \cdot 1 - 1)2^{K-n}, (2 \cdot 2 - 1)2^{K-n}, \cdots, (2 \cdot 2^{n-1} - 1)2^{K-n} \\ 0, & \text{otherwise.} \end{cases} \tag{120}$$

From (120), it can be shown that (110) is expressed as

$$d\big(S_j, S_{j+1}\big) = \sum_{n=1}^{K} f_n(j) d_{M_n} \quad \text{for } 1 \leq j \leq 2^K - 1. \tag{121}$$

C. BER of the MSB for hierarchical 2^K PAM

Fig. 9 depicts a hierarchical 2^K PAM constellation with the bit mapping vector $\mathbf{g}_1$ for the MSB given by (107). The system model for hierarchical 2^K PAM is shown in Fig. 10. The

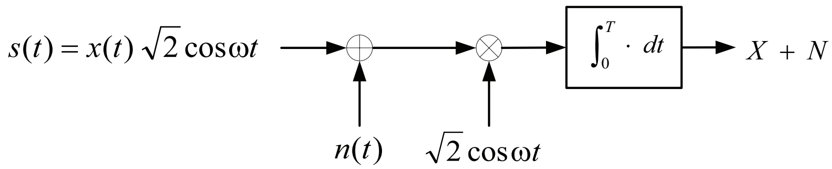

Fig. 10. System model for hierarchical PAM.

transmitted signal is given by

$$\begin{aligned} s(t) &= x(t)\sqrt{2}\cos\omega t \\ &= \operatorname{sgn}(i - 2^{K-1} - 0.5)d\big(0, S_i\big)P_T(t)\sqrt{2}\cos\omega t \quad \text{for } 1 \le i \le 2^K \end{aligned} \tag{122}$$

where sgn(·) denotes the sign of the real number, $d\big(0, S_i\big)$ is the Euclidian distance between the origin and ith signal point S_i ($1 \le i \le 2^K$), and $P_T(t)$ is the transmit pulse defined as

$$P_T(t) = \begin{cases} 1, & 0 \le t \le T \\ 0, & \text{elsewhere} \end{cases} \tag{123}$$

where T is the symbol duration. $n(t)$ is zero-mean additive white Gaussian noise having a power spectral density of $N_0/2$. At the receiver, the decision statistic is given by

$$X = \operatorname{sgn}(i - 2^{K-1} - 0.5)d\big(0, S_i\big)T \quad \text{and} \quad N = \int_0^T n(t)\sqrt{2}\cos\omega t\ dt \tag{124}$$

where the standard deviation of N is $\sqrt{N_0 T/2}$. From Fig. 9, since the decision boundary for bits 0 and 1 is the origin, the probability of correct decision for a signal point assigned for bit 1, S_i ($i > 2^{K-1} + 1$), is given by

$$P_{c,S_i} = \Pr\Big[0 \le d\big(0, S_i\big)T + N < \infty\Big] = 1 - Q\left(\frac{d\big(0, S_i\big)T}{\sqrt{N_0 T/2}}\right) = 1 - Q\left(d\big(0, S_i\big)\sqrt{\frac{2T}{N_0}}\right). \tag{125}$$

From (125), the probability of correct decision for the MSB is given by

$$P_c = \frac{1}{2^{K-1}} \sum_{i=2^{K-1}+1}^{2^K} P_{c,S_i} = 1 - \frac{1}{2^{K-1}} \sum_{i=2^{K-1}+1}^{2^K} Q\left(d\big(0, S_i\big)\sqrt{\frac{2T}{N_0}}\right) \tag{126}$$

and the BER for the MSB, P_{M_1}, is given by

$$P_{M_1} = 1 - P_c = \frac{1}{2^{K-1}} \sum_{i=2^{K-1}+1}^{2^K} Q\left(d\big(0, S_i\big)\sqrt{\frac{2T}{N_0}}\right). \tag{127}$$

From (110), for $n = 1$, we have

$$d\big(S_{(2i-1)2^{K-1}}, S_{(2i-1)2^{K-1}+1}\big) = d_{M_1} \quad \text{for } i = 1 \ \Leftrightarrow \ d\big(S_{2^{K-1}}, S_{2^{K-1}+1}\big) = d_{M_1}. \tag{128}$$

Since the hierarchical PAM constellation is symmetric with respect to the origin, from (128), we have

$$d\big(0, S_{2^{K-1}+1}\big) = \frac{1}{2} d\big(S_{2^{K-1}}, S_{2^{K-1}+1}\big) = \frac{d_{M_1}}{2}. \tag{129}$$

For $i \geq 2^{K-1}+2$, $d\big(0, S_i\big)$ can be expressed as

$$d\big(0, S_i\big) = d\big(0, S_{2^{K-1}+1}\big) + \sum_{j=2^{K-1}+1}^{i-1} d\big(S_j, S_{j+1}\big) = \frac{d_{M_1}}{2} + \sum_{j=2^{K-1}+1}^{i-1} d\big(S_j, S_{j+1}\big) \tag{130}$$

where the second equality follows from (129). From (129) and (130), the BER of the MSB, given by (127), can be rewritten as

$$P_{M_1} = \frac{1}{2^{K-1}} Q\left(\frac{d_{M_1}}{2}\sqrt{\frac{2T}{N_0}}\right) + \frac{1}{2^{K-1}} \sum_{i=2^{K-1}+2}^{2^K} Q\left(\left(\frac{d_{M_1}}{2} + \sum_{j=2^{K-1}+1}^{i-1} d\big(S_j, S_{j+1}\big)\right)\sqrt{\frac{2T}{N_0}}\right). \tag{131}$$

From (121), $\sum_{j=2^{K-1}+1}^{i-1} d\big(S_j, S_{j+1}\big)$ in (131) can be rewritten as

$$\sum_{j=2^{K-1}+1}^{i-1} d\big(S_j, S_{j+1}\big) = \sum_{j=2^{K-1}+1}^{i-1} \sum_{n=1}^{K} f_n(j) d_{M_n} = \sum_{n=1}^{K} d_{M_n} \sum_{j=2^{K-1}+1}^{i-1} f_n(j). \tag{132}$$

From (120), it can be shown that $\sum_{j=1}^{l} f_n(j)$ is expressed as

$$\sum_{j=1}^{l} f_n(j) = \left\lfloor \frac{l + 2^{K-n}}{2^{K-n+1}} \right\rfloor \quad \text{for } 1 \leq l \leq 2^K - 1 \text{ and } 1 \leq n \leq K. \tag{133}$$

From (133), (132) can be rewritten as

$$\sum_{j=2^{K-1}+1}^{i-1} d\big(S_j, S_{j+1}\big) = \sum_{n=1}^{K} d_{M_n} \left(\left\lfloor \frac{i - 1 + 2^{K-n}}{2^{K-n+1}} \right\rfloor - \left\lfloor \frac{2^{K-1} + 2^{K-n}}{2^{K-n+1}} \right\rfloor\right). \tag{134}$$

From (134), the second term of P_{M_1} given by (131) can be expressed as

$$\frac{1}{2^{K-1}} \sum_{i=2^{K-1}+2}^{2^K} Q\left(\left(\frac{d_{M_1}}{2} + \sum_{n=1}^{K} d_{M_n} \left\{\left\lfloor \frac{i - 1 + 2^{K-n}}{2^{K-n+1}} \right\rfloor - \left\lfloor \frac{2^{K-1} + 2^{K-n}}{2^{K-n+1}} \right\rfloor\right\}\right)\sqrt{\frac{2T}{N_0}}\right). \tag{135}$$

Let $p = i - 2^{K-1} - 1$. Then (135) can be rewritten as

$$\frac{1}{2^{K-1}} \sum_{p=1}^{2^{K-1}-1} Q\left(\left(\frac{d_{M_1}}{2} + \sum_{n=1}^{K} d_{M_n} \left\{ \left\lfloor \frac{p + 2^{K-n}}{2^{K-n+1}} + 2^{n-2} \right\rfloor - \left\lfloor 2^{n-2} + 2^{-1} \right\rfloor \right\} \right) \sqrt{\frac{2T}{N_0}}\right). \quad (136)$$

For $n \geq 2$, we have

$$\left\lfloor \frac{p + 2^{K-n}}{2^{K-n+1}} + 2^{n-2} \right\rfloor = \left\lfloor \frac{p + 2^{K-n}}{2^{K-n+1}} \right\rfloor + 2^{n-2} \quad \text{and} \quad \left\lfloor 2^{n-2} + 2^{-1} \right\rfloor = 2^{n-2}. \quad (137)$$

For $n = 1$, we have

$$\left\lfloor \frac{p + 2^{K-n}}{2^{K-n+1}} + 2^{n-2} \right\rfloor = \left\lfloor \frac{p}{2^K} + 1 \right\rfloor = 1 \quad \text{and} \quad \left\lfloor 2^{n-2} + 2^{-1} \right\rfloor = 1 \quad (138)$$

where the second equality of the first expression follows from $1 \leq p \leq 2^{K-1} - 1$ in (136). From (137) and (138), the second term of P_{M_1}, given by (136), can be rewritten as

$$\frac{1}{2^{K-1}} \sum_{p=1}^{2^{K-1}-1} Q\left(\left(\frac{d_{M_1}}{2} + \sum_{n=2}^{K} d_{M_n} \left\lfloor \frac{p + 2^{K-n}}{2^{K-n+1}} \right\rfloor \right) \sqrt{\frac{2T}{N_0}}\right). \quad (139)$$

Since $\sum_{n=2}^{K} d_{M_n} \left\lfloor \frac{p+2^{K-n}}{2^{K-n+1}} \right\rfloor = 0$ for $p = 0$, from (139), the BER of the MSB given by (131) can be expressed as

$$P_{M_1} = \frac{1}{2^{K-1}} \sum_{p=0}^{2^{K-1}-1} Q\left(\left(\frac{d_{M_1}}{2} + \sum_{n=2}^{K} d_{M_n} \left\lfloor \frac{p + 2^{K-n}}{2^{K-n+1}} \right\rfloor \right) \sqrt{\frac{2T}{N_0}}\right). \quad (140)$$

Note that (140) is the exact BER expression for the MSB of hierarchical 2^K PAM.

D. BER of the n_0th MSB $(2 \leq n_0 \leq K-1)$ for hierarchical 2^K PAM

D-1. Classification of 2^K signal points into 2^{n_0-1} mutually exclusive groups

We first find every pair of adjacent signal points which are separated by a Euclidian distance greater than $d_{M_{n_0}}$ (i.e., $d_{M_{n_0-1}}, d_{M_{n_0}-2}, \cdots, d_{M_1}$): For given n_0 in the range of $2 \leq n_0 \leq K-1$, let $n = n_0 - m$ $(1 \leq m \leq n_0 - 1)$ in (110). Then, we have

$$\begin{aligned} d\big(S_{(2i-1)2^{K-n_0+m}}, S_{(2i-1)2^{K-n_0+m}+1}\big) &= d_{M_{n_0-m}} \\ \text{for } 1 \leq i \leq 2^{n_0-m-1} &\text{ and } 1 \leq m \leq n_0 - 1. \end{aligned} \quad (141)$$

It can be shown that $\{(2i-1)2^{m-1} \mid 1 \leq i \leq 2^{n_0-m-1} \text{ and } 1 \leq m \leq n_0 - 1\}$ is identical to $\{j \mid 1 \leq j \leq 2^{n_0-1} - 1\}$. Hence, every pair of adjacent signal points which are separated by a Euclidian distance greater than $d_{M_{n_0}}$, given by (141), can be expressed as

$$S_{j \cdot 2^{K+1-n_0}}, \; S_{j \cdot 2^{K+1-n_0}+1} \qquad \text{for } 1 \leq j \leq 2^{n_0-1} - 1. \quad (142)$$

Next, we classify 2^K signal points into 2^{n_0-1} mutually exclusive groups such that the Euclidian distance between adjacent signal points of the same group is smaller than or equal to $d_{M_{n_0}}$. From (142), the signal points of the jth group can be derived as

$$S_{(j-1)2^{K+1-n_0}+1}, S_{(j-1)2^{K+1-n_0}+2}, \cdots, S_{j\cdot 2^{K+1-n_0}} \quad \text{for } 1 \le j \le 2^{n_0-1}. \tag{143}$$

We rewrite (110) in the following: For hierarchical 2^K PAM ($K \ge 2$), the Euclidian distance between adjacent signal points is given by

$$d\big(S_{(2i-1)2^{K-n}}, S_{(2i-1)2^{K-n}+1}\big) = d_{M_n} \quad \text{for } 1 \le i \le 2^{n-1} \text{ and } 1 \le n \le K. \tag{144}$$

From (143) and (144), it can be shown that the Euclidian distance between adjacent signal points of the jth group is given by

$$d\big(S_{(2i-1)2^{K-n}}, S_{(2i-1)2^{K-n}+1}\big) = d_{M_n}$$
$$\text{for } (j-1)2^{n-n_0}+1 \le i \le j\cdot 2^{n-n_0},\ n_0 \le n \le K, \text{ and } 1 \le j \le 2^{n_0-1}. \tag{145}$$

Let $p = i-(j-1)2^{n-n_0}$. Then, (145) can be rewritten as

$$d\big(S_{(2p-1)2^{K-n}+(j-1)2^{K+1-n_0}}, S_{(2p-1)2^{K-n}+(j-1)2^{K+1-n_0}+1}\big) = d_{M_n}$$
$$\text{for } 1 \le p \le 2^{n-n_0},\ n_0 \le n \le K, \text{ and } 1 \le j \le 2^{n_0-1}. \tag{146}$$

Notation change: Let $S_i^{(j)}$ denote $S_{(j-1)2^{K+1-n_0}+i}$ for convenience. Then, every pair of adjacent signal points which are separated by a Euclidian distance greater than $d_{M_{n_0}}$ (i.e., $d_{M_{n_0-1}}$, $d_{M_{n_0}-2}$, $\cdots$, d_{M_1}), given by (142), can be rewritten as

$$S^{(j)}_{2^{K+1-n_0}},\ S_1^{(j+1)} \quad \text{for } 1 \le j \le 2^{n_0-1}-1. \tag{147}$$

The signal points of the jth group, given by (143), can be expressed as

$$S_1^{(j)}, S_2^{(j)}, \cdots, S^{(j)}_{2^{K+1-n_0}} \quad \text{for } 1 \le j \le 2^{n_0-1}. \tag{148}$$

Lastly, the Euclidian distance between adjacent signal points of the jth group, given by (146), can be rewritten as

$$d\big(S^{(j)}_{(2p-1)2^{K-n}}, S^{(j)}_{(2p-1)2^{K-n}+1}\big) = d_{M_n}$$
$$\text{for } 1 \le p \le 2^{n-n_0},\ n_0 \le n \le K, \text{ and } 1 \le j \le 2^{n_0-1}. \tag{149}$$

D-2. Probability of correct decision for signal points of the jth group

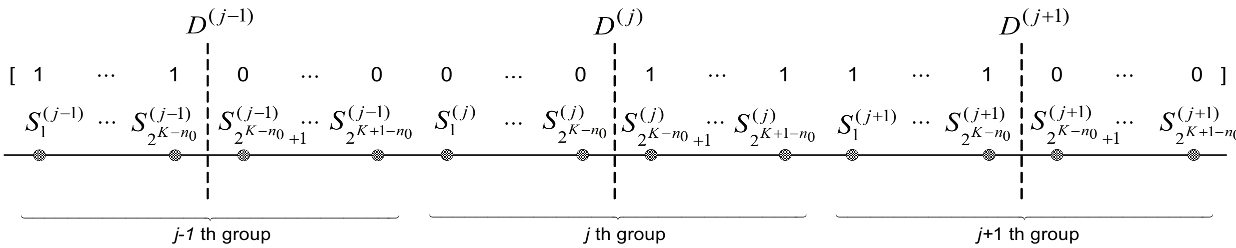

Fig. 11. The $j-1$, j and $j+1$th groups with the bit mapping vector for $j =$ odd.

From (148), for $2 \le j \le 2^{n_0-1} - 1$, the signal points of the $j-1$, j, and $j+1$th groups are given by

$$\underbrace{S_1^{(j-1)}, S_2^{(j-1)}, \cdots, S_{2^{K+1-n_0}}^{(j-1)}}_{j-1\text{th group}}, \underbrace{S_1^{(j)}, S_2^{(j)}, \cdots, S_{2^{K+1-n_0}}^{(j)}}_{j\text{th group}}, \underbrace{S_1^{(j+1)}, S_2^{(j+1)}, \cdots, S_{2^{K+1-n_0}}^{(j+1)}}_{j+1\text{th group}}. \tag{150}$$

From (107), the bit mapping vector for the n_0th MSB ($2 \le n_0 \le K-1$) of the $j-1$, j and $j+1$th groups is derived as

$$\begin{cases} \left[\mathbf{0}_{2^{K-n_0}} \ \mathbf{1}_{2^{K-n_0}} \ \mathbf{1}_{2^{K-n_0}} \ \mathbf{0}_{2^{K-n_0}} \ \mathbf{0}_{2^{K-n_0}} \ \mathbf{1}_{2^{K-n_0}}\right], & \text{for } j = \text{even} \\ \left[\mathbf{1}_{2^{K-n_0}} \ \mathbf{0}_{2^{K-n_0}} \ \mathbf{0}_{2^{K-n_0}} \ \mathbf{1}_{2^{K-n_0}} \ \mathbf{1}_{2^{K-n_0}} \ \mathbf{0}_{2^{K-n_0}}\right], & \text{for } j = \text{odd}. \end{cases} \tag{151}$$

From (150) and (151), $j-1$, j, and $j+1$th groups with the bit mapping vector for $j =$ odd are shown in Fig. 11, where $D^{(j-1)}$, $D^{(j)}$, and $D^{(j+1)}$ denote the decision boundaries for bits 0 and 1 in the $j-1$, j, and $j+1$th groups, respectively. In the following, we will derive the probability of correct decision for signal points of the jth group ($1 \le j \le 2^{n_0-1}$):

i) Signal points assigned for bit 0 when j is odd in the range of $2 \le j \le 2^{n_0-1} - 1$

We here assume that for $S_i^{(j)}$ ($1 \le i \le 2^{K-n_0}$), a signal point of the jth group which is assigned for bit 0, the probability of correct decision can be calculated without considering the other groups except for the $j-1$, j, and $j+1$ th groups (we will later show that the assumption is correct if the SNR condition of this lemma is satisfied). Fig. 12 shows the correct decision area for $S_i^{(j)}$ ($1 \le i \le 2^{K-n_0}$) under the above assumption. From Fig. 12, it follows that the probability of correct decision for $S_i^{(j)}$ ($1 \le i \le 2^{K-n_0}$) based on the system model depicted in

Fig. 10 is given by

$$\begin{aligned}P_c^{\mathrm{bit0}} &= \Pr\Big[-d\big(D^{(j-1)},S_i^{(j)}\big)T<N<d\big(S_i^{(j)},D^{(j)}\big)T\Big]\\&\quad+\Pr\Big[d\big(S_i^{(j)},D^{(j+1)}\big)T<N<d\big(S_i^{(j)},S_{2^{K+1-n_0}}^{(j+1)}\big)T\Big]\\&=\Pr\Big[-\Big(d\big(D^{(j-1)},S_1^{(j)}\big)+d\big(S_1^{(j)},S_i^{(j)}\big)\Big)T<N<\Big(d\big(S_i^{(j)},S_{2^{K-n_0}}^{(j)}\big)+d\big(S_{2^{K-n_0}}^{(j)},D^{(j)}\big)\Big)T\Big]\\&\quad+\Pr\Big[\Big(d\big(S_i^{(j)},S_{2^{K-n_0}}^{(j)}\big)+d\big(S_{2^{K-n_0}}^{(j)},D^{(j+1)}\big)\Big)T<N<\big(S_i^{(j)},S_{2^{K+1-n_0}}^{(j+1)}\big)T\Big]\end{aligned}\tag{152}$$

where the first and second terms follow from the correct decision areas #1 and #2 shown in Fig. 12, respectively. Eq. (152) can be rewritten as

$$\begin{aligned}P_c^{\mathrm{bit0}} &= 1-\Pr\Big[N>\Big(d\big(S_i^{(j)},S_{2^{K-n_0}}^{(j)}\big)+d\big(S_{2^{K-n_0}}^{(j)},D^{(j)}\big)\Big)T\Big]\\&\quad-\Pr\Big[N>\Big(d\big(D^{(j-1)},S_1^{(j)}\big)+d\big(S_1^{(j)},S_i^{(j)}\big)\Big)T\Big]\\&\quad+\Pr\Big[\Big(d\big(S_i^{(j)},S_{2^{K-n_0}}^{(j)}\big)+d\big(S_{2^{K-n_0}}^{(j)},D^{(j+1)}\big)\Big)T<N<\big(S_i^{(j)},S_{2^{K+1-n_0}}^{(j+1)}\big)T\Big].\end{aligned}\tag{153}$$

From Fig. 12, $d\big(D^{(j-1)},S_1^{(j)}\big)$ in the second term of (153) can be expressed as

$$d\big(D^{(j-1)},S_1^{(j)}\big)=d\big(D^{(j-1)},S_{2^{K-n_0}+1}^{(j-1)}\big)+d\big(S_{2^{K-n_0}+1}^{(j-1)},S_{2^{K+1-n_0}}^{(j-1)}\big)+d\big(S_{2^{K+1-n_0}}^{(j-1)},S_1^{(j)}\big).\tag{154}$$

From (149), for $n=n_0$, we have

$$d\big(S_{2^{K-n_0}}^{(j)},S_{2^{K-n_0}+1}^{(j)}\big)=d_{M_{n_0}}\quad\text{for } 1\le j\le 2^{n_0-1}.\tag{155}$$

From the fact that $d\big(D^{(j-1)},S_{2^{K-n_0}+1}^{(j-1)}\big)=\frac{1}{2}d\big(S_{2^{K-n_0}}^{(j-1)},S_{2^{K-n_0}+1}^{(j-1)}\big)$ and (155), (154) can be rewritten as

$$d\big(D^{(j-1)},S_1^{(j)}\big)=\frac{1}{2}d_{M_{n_0}}+d\big(S_{2^{K-n_0}+1}^{(j-1)},S_{2^{K+1-n_0}}^{(j-1)}\big)+d\big(S_{2^{K+1-n_0}}^{(j-1)},S_1^{(j)}\big).\tag{156}$$

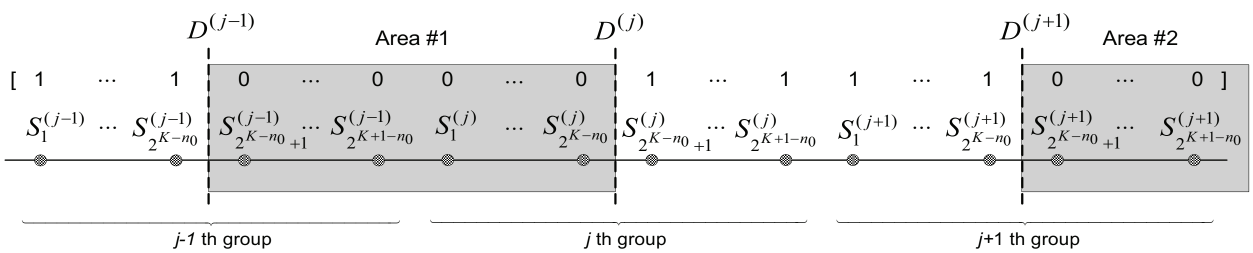

Fig. 12. The correct decision area for $S_i^{(j)}$ ($1\le i\le 2^{K-n_0}$) when $j=\text{odd}$.

From (147), we have

$$d\big(S^{(j)}_{2^{K+1-n_0}}, S^{(j+1)}_1\big) \geq d_{M_{n_0-1}} \quad \text{for } 1 \leq j \leq 2^{n_0-1}-1. \tag{157}$$

From (157), $d\big(D^{(j-1)}, S^{(j)}_1\big)$, given by (156), satisfies

$$d\big(D^{(j-1)}, S^{(j)}_1\big) \geq \frac{1}{2}d_{M_{n_0}} + d_{M_{n_0-1}} + d\big(S^{(j-1)}_{2^{K-n_0}+1}, S^{(j-1)}_{2^{K+1-n_0}}\big). \tag{158}$$

Since $d\big(S^{(j-1)}_{2^{K-n_0}+1}, S^{(j-1)}_{2^{K+1-n_0}}\big) > 0$ for $n_0 \leq K-1$, we have

$$d\big(D^{(j-1)}, S^{(j)}_1\big) > \frac{1}{2}d_{M_{n_0}} + d_{M_{n_0-1}}. \tag{159}$$

Likewise, from Fig. 12, $d\big(S^{(j)}_{2^{K-n_0}}, D^{(j+1)}\big)$ in the third term of (153) can be expressed as

$$\begin{aligned} d\big(S^{(j)}_{2^{K-n_0}}, D^{(j+1)}\big) = {} & d\big(S^{(j)}_{2^{K-n_0}}, S^{(j)}_{2^{K-n_0}+1}\big) + d\big(S^{(j)}_{2^{K-n_0}+1}, S^{(j)}_{2^{K+1-n_0}}\big) + d\big(S^{(j)}_{2^{K+1-n_0}}, S^{(j+1)}_1\big) \\ & + d\big(S^{(j+1)}_1, S^{(j+1)}_{2^{K-n_0}}\big) + d\big(S^{(j+1)}_{2^{K-n_0}}, D^{(j+1)}\big). \end{aligned} \tag{160}$$

Since $d\big(S^{(j+1)}_{2^{K-n_0}}, D^{(j+1)}\big) = \frac{1}{2}d\big(S^{(j+1)}_{2^{K-n_0}}, S^{(j+1)}_{2^{K-n_0}+1}\big)$ and from (155), (160) can be rewritten as

$$\begin{aligned} d\big(S^{(j)}_{2^{K-n_0}}, D^{(j+1)}\big) = {} & \frac{3}{2}d_{M_{n_0}} + d\big(S^{(j)}_{2^{K-n_0}+1}, S^{(j)}_{2^{K+1-n_0}}\big) + d\big(S^{(j)}_{2^{K+1-n_0}}, S^{(j+1)}_1\big) \\ & + d\big(S^{(j+1)}_1, S^{(j+1)}_{2^{K-n_0}}\big). \end{aligned} \tag{161}$$

From (157), $d\big(S^{(j)}_{2^{K-n_0}}, D^{(j+1)}\big)$ satisfies

$$d\big(S^{(j)}_{2^{K-n_0}}, D^{(j+1)}\big) = \frac{3}{2}d_{M_{n_0}} + d_{M_{n_0-1}} + d\big(S^{(j)}_{2^{K-n_0}+1}, S^{(j)}_{2^{K+1-n_0}}\big) + d\big(S^{(j+1)}_1, S^{(j+1)}_{2^{K-n_0}}\big). \tag{162}$$

We have $d\big(S^{(j)}_{2^{K-n_0}+1}, S^{(j)}_{2^{K+1-n_0}}\big) > 0$ and $d\big(S^{(j+1)}_1, S^{(j+1)}_{2^{K-n_0}}\big) > 0$ for $n_0 \leq K-1$. Hence, $d\big(S^{(j)}_{2^{K-n_0}}, D^{(j+1)}\big)$ satisfies

$$d\big(S^{(j)}_{2^{K-n_0}}, D^{(j+1)}\big) > \frac{3}{2}d_{M_{n_0}} + d_{M_{n_0-1}}. \tag{163}$$

From (159) and (163), it follows that the second and third terms of P_c^{bit0}, given by (153), are insignificant when the condition of this lemma is satisfied. Since $S^{(j)}_i$, $S^{(j)}_{2^{K-n_0}}$, and $D^{(j)}$ belong to the jth group, $d\big(S^{(j)}_i, S^{(j)}_{2^{K-n_0}}\big) + d\big(S^{(j)}_{2^{K-n_0}}, D^{(j)}\big)$ in the first term of P_c^{bit0} is the combination of $d_{M_{n_0}}, d_{M_{n_0+1}}, \cdots, d_{M_K}$ from (149), and thus the first term is not affected by the condition of this lemma. Hence, if the condition of this lemma is satisfied, P_c^{bit0}, given by (153), becomes

$$P_c^{\text{bit0}} \approx 1 - \Pr\Big[N > \Big(d\big(S^{(j)}_i, S^{(j)}_{2^{K-n_0}}\big) + d\big(S^{(j)}_{2^{K-n_0}}, D^{(j)}\big)\Big)T\Big], \tag{164}$$

which is identical to the probability of correct decision calculated only by considering 2^{K+1-n_0} signal points of the isolated jth group. Since the $j-1$ and $j+1$th groups have no effect on the correct decision probability for signal points of the jth group due to the condition of this lemma, the other groups (i.e., $1, \cdots, j-2, j+2, \cdots, 2^{n_0-1}$th groups), which are separated by larger Euclidian distances from the jth group than are the $j-1$ and $j+1$th groups, also have no effect. Hence, the assumption above (152) is correct.

ii) Signal points assigned for bit 1 when j is odd in the range of $2 \le j \le 2^{n_0-1}-1$

It can be shown that the probability of correct decision for $S_i^{(j)}$ ($2^{K-n_0}+1 \le i \le 2^{K+1-n_0}$) based on the system model depicted in Fig. 10 is given by

$$\begin{aligned} P_c^{\mathrm{bit1}} = 1 &- \Pr\Big[N > \Big(d\big(D^{(j)}, S^{(j)}_{2^{K-n_0}+1}\big) + d\big(S^{(j)}_{2^{K-n_0}+1}, S_i^{(j)}\big)\Big)\, T\Big] \\ &-\Pr\Big[N > \Big(d\big(S_i^{(j)}, S^{(j)}_{2^{K+1-n_0}}\big) + d\big(S^{(j)}_{2^{K+1-n_0}}, D^{(j+1)}\big)\Big)\, T\Big] \\ &+\Pr\Big[\, \Big(d\big(D^{(j-1)}, S^{(j)}_{2^{K-n_0}+1}\big) + d\big(S^{(j)}_{2^{K-n_0}+1}, S_i^{(j)}\big)\Big)\, T < N < d\big(S_1^{(j-1)}, S_i^{(j)}\big) T\Big], \end{aligned} \tag{165}$$

where $d\big(S^{(j)}_{2^{K+1-n_0}}, D^{(j+1)}\big)$in the second term of (165) satisfies [1]

$$d\big(S^{(j)}_{2^{K+1-n_0}}, D^{(j+1)}\big) > \frac{1}{2} d_{M_{n_0}} + d_{M_{n_0-1}}, \tag{166}$$

and $d\big(D^{(j-1)}, S^{(j)}_{2^{K-n_0}+1}\big)$ in the third term of (165) satisfies

$$d\big(D^{(j-1)}, S^{(j)}_{2^{K-n_0}+1}\big) > \frac{3}{2} d_{M_{n_0}} + d_{M_{n_0-1}}. \tag{167}$$

From (165)–(167), if the condition of this lemma is satisfied, P_c^{bit1}, given by (165), becomes

$$P_c^{\mathrm{bit1}} \approx 1 - \Pr\Big[N > \Big(d\big(D^{(j)}, S^{(j)}_{2^{K-n_0}+1}\big) + d\big(S^{(j)}_{2^{K-n_0}+1}, S_i^{(j)}\big)\Big)\, T\Big], \tag{168}$$

which is identical to the probability of correct decision calculated only by considering 2^{K+1-n_0} signal points of the isolated jth group.

iii) Signal points assigned for bit 0/1 when j is even in the range of $2 \le j \le 2^{n_0-1}-1$

From (151), the bit mapping vector for $j =$ even is just the complement of that for $j =$ odd. Hence, for $j =$ even, P_c^{bit0} and P_c^{bit1} are given by (165) and (153), respectively, and the results of i) and ii) hold for the case j is even.

iv) Signal points assigned for bit 0/1 when $j = 1$ (odd)

[1] Since the analysis of ii) is similar to that of i), we omit the detailed steps.

From Fig. 12, it follows that P_c^{bit0} for $j=1$ is given by

$$P_c^{\text{bit0}} = 1 - \Pr\Big[N > \Big(d\big(S_i^{(1)}, S_{2^{K-n_0}}^{(1)}\big) + d\big(S_{2^{K-n_0}}^{(1)}, D^{(1)}\big)\Big)\,T\Big] \\ +\Pr\Big[\,\Big(d\big(S_i^{(1)}, S_{2^{K-n_0}}^{(1)}\big) + d\big(S_{2^{K-n_0}}^{(1)}, D^{(2)}\big)\Big)\,T < N < d\big(S_i^{(1)}, S_{2^{K+1-n_0}}^{(2)}\big)T\Big]. \quad (169)$$

The only difference between (153) and (169) is that (169) does not have the second term of (153), and thus the result of i) holds for the case $j=1$. In a similar way, it can be shown that for bit 1, the result of ii) holds for the case $j=1$.

v) Signal points assigned for bit 0/1 when $j = 2^{n_0-1}$ (even)

In a similar way to iv), it can be shown that the result of iii) holds for the case $j = 2^{n_0-1}$.

From i)–v), it is seen that if the SNR condition of this lemma is satisfied, the BER of the n_0th MSB can be calculated only by considering 2^{K+1-n_0} signal points of the isolated jth group given by (148).

D-3. BER of the n_0th MSB $(2 \le n_0 \le K-1)$ for the isolated jth group

We derive the BER of the n_0th MSB for the isolated jth group of 2^K PAM from that of the MSB for 2^{K+1-n_0} PAM.

i) For hierarchical 2^{K+1-n_0} PAM, from (144), the Euclidian distance between adjacent signal points is given by

$$d\big(S_{(2i-1)2^{K+1-n_0-n}}, S_{(2i-1)2^{K+1-n_0-n}+1}\big) = d_{M_n} \\ \text{for } 1 \le i \le 2^{n-1} \text{ and } 1 \le n \le K+1-n_0. \quad (170)$$

Let $r = n + n_0 - 1$ and $p = i$. Then, (170) can be rewritten as

$$d\big(S_{(2p-1)2^{K-r}}, S_{(2p-1)2^{K-r}+1}\big) = d_{M_{r+1-n_0}} \quad \text{for } 1 \le p \le 2^{r-n_0} \text{ and } n_0 \le r \le K. \quad (171)$$

From (149) and (171), it is seen that, if $d_{M_{r+1-n_0}}$ in (171) is set equal to d_{M_r}, the Euclidian distance between adjacent signal points for 2^{K+1-n_0} PAM is the same as that for the jth group of 2^K PAM.

ii) For hierarchical 2^{K+1-n_0} PAM, from (107), the bit mapping vector for the MSB is given by

$$\mathbf{g}_1 = \big[\mathbf{0}_{2^{(K+1-n_0)-1}}\ \mathbf{1}_{2^{(K+1-n_0)-1}}\big] = \big[\mathbf{0}_{2^{K-n_0}}\ \mathbf{1}_{2^{K-n_0}}\big]. \quad (172)$$

For the jth group of hierarchical 2^K PAM, from (151), the bit mapping vector for the n_0th MSB is given by

$$\begin{cases} \left[\mathbf{1}_{2^{K-n_0}}\ \mathbf{0}_{2^{K-n_0}}\right], & \text{for } j = \text{even} \\ \left[\mathbf{0}_{2^{K-n_0}}\ \mathbf{1}_{2^{K-n_0}}\right], & \text{for } j = \text{odd}. \end{cases} \tag{173}$$

It is seen that (173) is the same as or the complement of (172).

From i) and ii), it follows that the BER of the MSB for 2^{K+1-n_0} PAM is the same as that of the n_0th MSB for the isolated jth group of 2^K PAM, if $d_{M_{r+1-n_0}}$ for 2^{K+1-n_0} PAM is set equal to d_{M_r} (i.e., d_{M_x} is set equal to $d_{M_{x+n_0-1}}$). From (140), the BER of the MSB for hierarchical 2^{K+1-n_0} PAM ($2 \leq n_0 \leq K-1$) is derived as

$$\frac{1}{2^{K-n_0}} \sum_{p=0}^{2^{K-n_0}-1} Q\left(\left(\frac{d_{M_1}}{2} + \sum_{n=2}^{K+1-n_0} \left\lfloor \frac{p+2^{K+1-n_0-n}}{2^{K+2-n_0-n}} \right\rfloor d_{M_n}\right)\sqrt{\frac{2T}{N_0}}\right). \tag{174}$$

Let $r = n_0 - 1 + n$. Then (174) can be rewritten as

$$\frac{1}{2^{K-n_0}} \sum_{p=0}^{2^{K-n_0}-1} Q\left(\left(\frac{d_{M_1}}{2} + \sum_{r=n_0+1}^{K} \left\lfloor \frac{p+2^{K-r}}{2^{K-r+1}} \right\rfloor d_{M_{r-n_0+1}}\right)\sqrt{\frac{2T}{N_0}}\right). \tag{175}$$

As stated above (174), by setting d_{M_x} equal to $d_{M_{x+n_0-1}}$ in (175), the BER for the n_0th MSB ($2 \leq n_0 \leq K-1$) of the isolated jth group can be derived as

$$\frac{1}{2^{K-n_0}} \sum_{p=0}^{2^{K-n_0}-1} Q\left(\left(\frac{d_{M_{n_0}}}{2} + \sum_{r=n_0+1}^{K} \left\lfloor \frac{p+2^{K-r}}{2^{K-r+1}} \right\rfloor d_{M_r}\right)\sqrt{\frac{2T}{N_0}}\right). \tag{176}$$

Note that the BER expression for the n_0th MSB ($2 \leq n_0 \leq K-1$) of hierarchial 2^K PAM, given by (176), holds if the condition of this lemma is satisfied.

E. BER of the Kth MSB (or LSB) for hierarchical 2^K PAM

For the Kth MSB (or LSB), we define the signal points of the jth group as

$$S_1^{(j)}, S_2^{(j)}, S_3^{(j)}, S_4^{(j)} \quad \text{for } 1 \leq j \leq 2^{K-2}, \tag{177}$$

which is identical to (148) with $n_0 = K-1$. If we let $n_0 = K-1$ in (147), every pair of adjacent signal points which are separated by Euclidian distances greater than $d_{M_{K-1}}$ is given by

$$S_4^{(j)},\ S_1^{(j+1)} \quad \text{for } 1 \leq j \leq 2^{K-2}-1. \tag{178}$$

Also let $n_0 = K-1$ in (149). Then, for $1 \leq j \leq 2^{K-2}$, the Euclidian distance between adjacent signal points of the jth group can be derived as

$$d\big(S_2^{(j)}, S_3^{(j)}\big) = d_{M_{K-1}} \quad \text{and} \quad d\big(S_1^{(j)}, S_2^{(j)}\big) = d\big(S_3^{(j)}, S_4^{(j)}\big) = d_{M_K}. \tag{179}$$

From (107) and (177) – (179), it can be shown that if the condition of this lemma is satisfied, the BER of the Kth MSB becomes [2]

$$Q\left(\frac{d_{M_K}}{2}\sqrt{\frac{2T}{N_0}}\right) + \frac{1}{2}Q\left(\left(d_{M_{K-1}} + \frac{d_{M_K}}{2}\right)\sqrt{\frac{2T}{N_0}}\right). \tag{180}$$

From (140), (176), and (180), the BER of the n_0th MSB ($1 \leq n_0 \leq K$) for hierarchical 2^K PAM can be expressed as

$$\begin{cases} \sum_{p=0}^{2^{K-n_0}-1} \frac{1}{2^{K-n_0}} Q\left(\left(\frac{d_{M_{n_0}}}{2} + \sum_{r=n_0+1}^{K} \left\lfloor \frac{p+2^{K-r}}{2^{K-r+1}} \right\rfloor d_{M_r}\right)\sqrt{\frac{2T}{N_0}}\right), & \text{for } 1 \leq n_0 \leq K-1 \\ Q\left(\frac{d_{M_K}}{2}\sqrt{\frac{2T}{N_0}}\right) + \frac{1}{2}Q\left(\left(d_{M_{K-1}} + \frac{d_{M_K}}{2}\right)\sqrt{\frac{2T}{N_0}}\right), & \text{for } n_0 = K. \end{cases} \tag{181}$$

Note that (181) is the exact BER expression for the MSB, but for $2 \leq n_0 \leq K$ th MSB, (181) holds if the condition of this lemma is satisfied. Lastly, it can be shown that the BER of the inphase or quadrature components for hierarchical 2^{2K} QAM is the same as that for hierarchical 2^K PAM. For hierarchical 2^{2K} QAM, let $E_s = P_{avg}T$ denote the average energy of the transmitted signal. Setting $2T/N_0 = 2E_s/N_0 P_{avg} = 2\gamma_s/P_{avg}$ in (181), (16) is derived.

APPENDIX B
NUMERICAL EVALUATION OF THE BER EXPRESSION (16)

Figs. 13 and 14 show the numerical evaluation of the BER expression given by (16) for hierarchical 64 and 256 QAM when the distance ratio is 1 or 2.

REFERENCES

[1] J. F. Hayes, "Adaptive feedback communications," *IEEE Trans. Commun. Technol.*, vol. 16, no. 1, pp. 29–34, Feb. 1968.

[2] W. T. Webb and R. Steele, "Variable rate QAM for mobile radio," *IEEE Trans. Commun.*, vol. 43, no. 7, pp. 2223–2230, July 1995.

[3] A. J. Goldsmith and P. P. Varaiya, "Capacity of fading channels with channel side information," *IEEE Trans. Inform. Theory*, vol. 43, no. 6, pp. 1986–1992, Nov. 1997.

[2] Since the analysis for the Kth MSB is similar to that for the n_0th MSB ($2 \leq n_0 \leq K-1$), we omit the detailed steps but they can be found in [32].

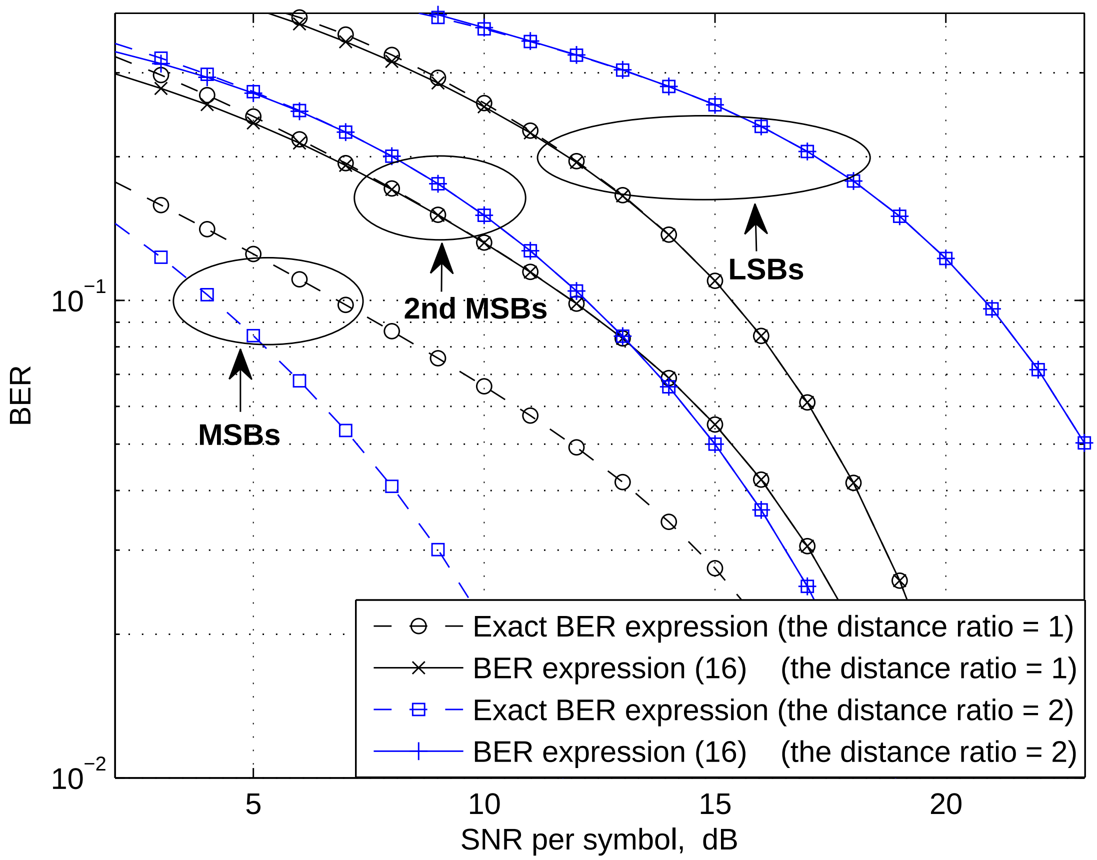

Fig. 13. BER for hierarchical 64 QAM: The distance ratio $d_{M_{n-1}}/d_{M_n}$ = 1 (i.e, the lower bound) or 2.

[4] T. Cover, "Broadcast channels," *IEEE Trans. Inform. Theory*, vol. 18, no. 1, pp. 2–14, Jan. 1972.

[5] P. P. Bergmans, "Random coding theorem for broadcast channels with degraded components," *IEEE Trans. Inform. Theory*, vol. 19, no. 2, pp. 197–207, Mar. 1973.

[6] E. C. van der Meulen, "A survey of multi-way channels in information theory: 1961–1976," *IEEE Trans. Inform. Theory*, vol. 23, pp. 1–37, Jan. 1977.

[7] K. Ramchandran, A. Ortega, K.M. Uz, and M.Vetterli, "Multiresolution broadcast for digital HDTV using joint source/channel coding," *IEEE J. Select. Areas Commun.*, vol. 11, no. 1, pp. 6–23, Jan. 1993.

[8] L.-F. Wei, "Coded modulation with unequal error protection," *IEEE Trans. Commun.*, vol. 41, no. 10, pp. 1439–1449, Oct. 1993.

[9] A. R. Calderbank and N. Seshadri, "Multilevel codes for unequal error protection," *IEEE Trans. Inform. Theory*, vol. 39, no. 4, pp. 1234–1248, July 1993.

[10] M. Morimoto, H. Harada, M. Okada, and S. Komaki, "A study on power assignment of hierarchical modulation schemes for digital broadcasting," *IEICE Trans. Commun.*, vol. E77-B, pp. 1495–1500, Dec. 1994.

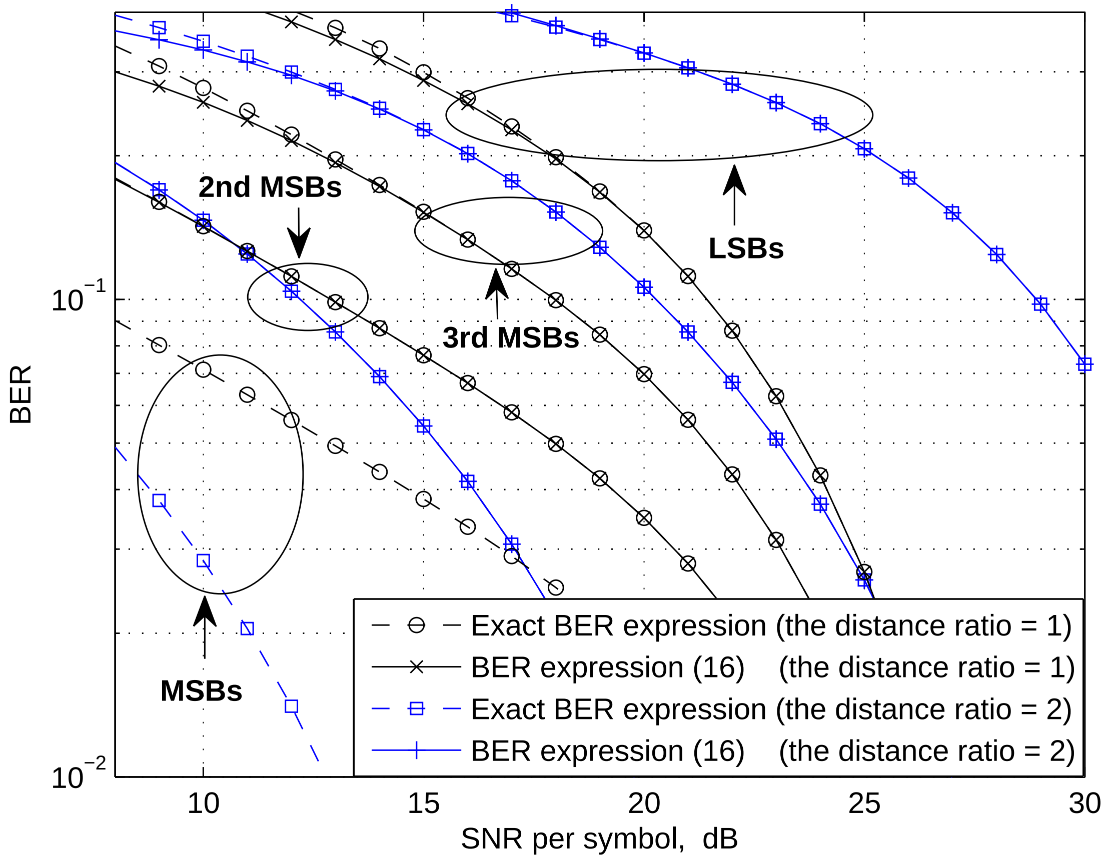

Fig. 14. BER for hierarchical 256 QAM: The distance ratio $d_{M_{n-1}}/d_{M_n}$ = 1 (i.e, the lower bound) or 2.

[11] C.-E. W. Sundberg, W. C. Wong, and R. Steele, "Logarithmic PCM weighted QAM transmission over Gaussian and Rayleigh fading channels," *Proc. Inst. Elec. Eng.*, vol. 134, pp. 557–570, Oct. 1987.

[12] "Digital Video Broadcasting (DVB); Framing Structure, Channel Coding and Modulation for Digital Terrestrial Television," ETSI EN 300 744 V1.5.1, Nov. 2004.

[13] C. Weck, "Coverage aspects of digital terrestrial television broadcasting," *EBU Techn. Rev.*, pp. 19–30, Winter 1996.

[14] A. Schertz and C.Weck, "Hierarchical modulation-the transmission of two independent DVB-T multiplexes on a single frequency," *EBU Techn. Rev.*, Apr. 2003.

[15] M. B. Pursley and J. M. Shea, "Nonuniform phase-shift-key modulation for multimedia multicast transmission in mobile wireless networks," *IEEE J. Select. Areas Commun.*, vol. 5, pp. 774–783, May 1999.

[16] M. B. Pursley and J. M. Shea, "Adaptive nonuniform phase-shift-key modulation for multimedia traffic in wireless networks," *IEEE J. Select. Areas Commun.*, vol. 18, pp. 1394–1407, Aug. 2000.

[17] B. Masnick and J. Wolf, "On linear unequal error protection codes," *IEEE Trans. Inform. Theory*, vol. 13, pp. 600–607, Oct. 1967.

[18] H. Suda and T. Miki, "An error protected 16 kb/s voice transmission for land mobile radio channel," *IEEE J. Select. Areas Commun.*, vol. 6, pp. 346–352, Feb. 1988.

[19] R. V. Cox, J. Hagenauer, N. Seshadri, and C.-E. W. Sundberg, "Subband speech coding and matched convolutional channel coding for mobile radio channels," *IEEE Trans. Signal Process.*, vol. 39, pp. 1717–1731, Aug. 1991.

[20] P. G. Sherwood and K. Zeger, "Progressive image coding for noisy channels," *IEEE Signal Process. Lett.*, vol. 4, no. 7, pp. 191–198, July 1997.

[21] P. Cosman, J. Rogers, P. G. Sherwood, and K. Zeger, "Combined forward error control and packetized zerotree wavelet encoding for transmission of images over varying channels," *IEEE Trans. Image Process.*, vol. 9, no. 6, pp. 132–140, June 2000.

[22] V. Stankovic, R. Hamzaoui, Y. Charfi, and Z. Xiong, "Real-time unequal error protection algorithms for progressive image transmission," *IEEE J. Select. Areas Commun.*, vol. 21, no. 10, Dec. 2003.

[23] Y. Pei and J. W. Modestino, "Cross-Layer Design for Video Transmission over Wireless Rician Slow-Fading Channels Using an Adaptive Multiresolution Modulation and Coding Scheme," *EURASIP J. Advances Signal Process.*, vol. 2007, article ID 86915.

[24] B. Barmada, M.M. Ghandi, E.V. Jones, and M. Ghanbari, "Prioritized transmission of data partitioned H.264 video with hierarchical QAM," *IEEE Signal Process. Lett.*, vol. 12, pp. 577–580, Aug. 2005.

[25] J. M. Shapiro, "Embedded image coding using zerotrees of wavelet coefficients," *IEEE Trans. Signal Process.*, vol. 41, no. 12, pp. 3445–3462, Dec. 1993.

[26] A. Said and W. A. Pearlman, "A new, fast, and efficient image codec based on set partitioning in hierarchical trees," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 6, no. 3, pp. 243–249, June 1996.

[27] D. Taubman and M. Marcellin, JPEG2000: Image Compression Fundamentals, Standards, and Practice. Norwell, MA: Kluwer, 2001.

[28] J. Reichel, H. Schwarz, and M. Wien (eds.), "Scalable Video Coding -Working Draft 1," Joint Video Team of ITU-T VCEG and ISO/IEC MPEG, Doc. JVT-N020, Hong Kong, CN, Jan. 2005.

[29] H. Schwarz, D. Marpe, and T. Wiegand, "Overview of the scalable video coding extension of H.264/AVC," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 17, no. 9, pp. 1103–1120, Sept. 2007.

[30] F. Wu, S. Li, and Y.-Q. Zhang, "A framework for efficient progressive fine granularity scalable video coding," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 11, no. 3, pp. 282–300, Mar. 2001.

[31] P. K. Vitthaladevuni and M.-S. Alouini, "A recursive algorithm for the exact BER computation of generalized hierarchical QAM constellations," *IEEE Trans. Inform. Theory*, vol. 49, pp. 297–307, Jan. 2003.

[32] S.-H. Chang, "Joint optimization of physical and application layers for wireless multimedia communications," Ph.D. dissertation, University of California, San Diego, 2010.

[33] J. G. Proakis, *Digital Communication*, 3rd ed. New York: McGraw-Hill, 1995.