

Mitigating Bit-Synchronization Errors in Huffman-Coding-Aided Index Modulation

Ye Liu and Justin P. Coon, *Senior Member, IEEE*

Abstract—When utilizing all subcarrier activation patterns (SAPs) in the combinatorial index modulation (IM) scheme, existing techniques either fail to realize the maximum throughput enhancement or fail to address the bit-level synchronization issue (or error propagation) due to the variable-length nature of the index-domain codebook. To address these drawbacks, we propose using Huffman coding for IM together with two constellation mappings for the bandpass modulation. The proposed scheme induces no signal overhead and is applicable when the index-domain codebook has the maximum encoding rate. Simulations demonstrate the effectiveness of the proposed scheme in mitigating bit-level synchronization errors.

I. INTRODUCTION

As a special case of permutation modulation [1], index modulation (IM) is a technique for multicarrier communication systems, e.g., orthogonal frequency-division multiplexing (OFDM), where the subcarrier activation patterns (SAPs) carry part of the information. By activating a fraction of all subcarriers, IM can improve the throughput of the multicarrier system and may reduce the peak-to-average power ratio (PAPR). Among many possible designs of mapping information to SAPs (e.g., [2], [3], [4]), the combinatorial approach in [4], commonly referred to as OFDM-IM, stands out due to its relatively high throughput and good error performance.

When encoding a fixed number of bits in the index domain, [4] is not always able to utilize all SAPs, which does not realize the full potential of IM in throughput enhancement and can lead to catastrophic errors at the decoder. To utilize all SAPs, [5] and [6] propose variable-length coding for the index domain messages, and the same technique is potentially applicable to systems other than OFDM [7]. However, it is not clear from [5] and [6] what is the *maximum encoding rate*, i.e., at most how many bits are encoded into an SAP on average, of such a variable-length code. Therefore, the maximum throughput gain of Huffman-coding-aided OFDM-IM versus conventional OFDM-IM remains to be revealed.

Moreover, bit-level synchronization errors¹ arise when the index codebook encodes bit sequences of variable lengths. The synchronization issue is, however, not addressed by [6]. The same issue is mitigated in [5] at the cost of reducing the throughput of an OFDM symbol, which limits the benefit of utilizing all SAPs to removing catastrophic errors only.

This work aims to mitigate bit-synchronization errors when enhancing the index-domain encoding rate via variable-length codes. Specifically,

- 1) We point out the properties of the binary trees that lead to index codebooks of maximum encoding rate, where we assume a maximally entropic information source in the index domain. We show that the design in [5] is equivalent to Huffman codebooks that maximize the encoding rate.
- 2) By observing that an index codebook of maximum encoding rate encodes bit sequences of at most two different lengths, we propose using two constellations for the bandpass signals to distinguish the lengths of the index domain messages. Such an approach has the benefit that no signal overhead is added to the system.

Simulation results show that the proposed constellation adaptation can harvest the throughput gain thanks to Huffman coding with improved bit error rate (BER) performance compared to OFDM-IM for certain system settings.

II. SYSTEM MODEL

Consider an OFDM system with N subcarriers realized by a size- N discrete Fourier transform and its inverse. To implement the combinatorial IM scheme [4], first, the N subcarriers are divided into G groups, each containing $n = N/G$ subcarriers. Then, k out of n subcarriers in each group are activated. Assume an M -ary constellation is applied to the activated subcarriers. Also, let p_1 denote the number of bits carried in the index domain, and let p_2 denote the number of bits carried by the constellation symbols mapped onto the activated subcarriers. The scheme in [4] constructs an index codebook between the index-domain message and the SAPs that gives a throughput of $p_1 + p_2$ bits per OFDM symbol, where $p_1 = G \cdot \lfloor \log_2 \binom{n}{k} \rfloor$, $p_2 = kG \cdot \log_2(M)$, and $\binom{n}{k}$ is the binomial coefficient. The floor function $\lfloor \cdot \rfloor$ appears in p_1 because we cannot encode a fraction of one bit in a digital system. Therefore, $D := \binom{n}{k} - 2^{\lfloor \log_2 \binom{n}{k} \rfloor}$ SAPs are left unused.

In the next section, we apply binary-tree based coding to utilize all $\binom{n}{k}$ SAPs. Such a technique has been briefly discussed by [5] and [6], but there is a lack of understanding on what is, and how to achieve, the maximum encoding rate of the technique, which is clarified in the next section.

III. BINARY-TREE INDEX CODEBOOK

In this section, we show that certain Huffman codebooks maximize the encoding rate of the index codebook.

A. IM through Huffman Coding

The ability to map variable-length bit sequences to symbols makes Huffman codes suitable for utilizing all $\binom{n}{k}$ SAPs in

The authors are with the Department of Engineering Science, University of Oxford, OX1 3PJ, United Kingdom (email: ye.liu, justin.coon@eng.ox.ac.uk).

This work has been supported by the Engineering and Physical Science Research Council (EPSRC) Impact Acceleration Account Award EP/R511742/1.

¹Also referred to as error propagation.

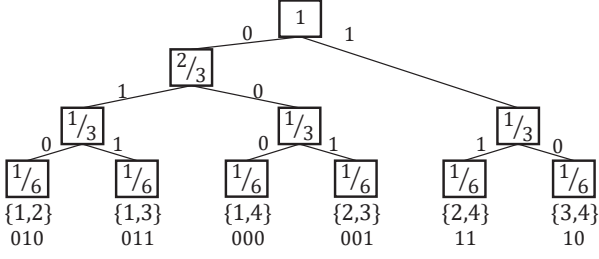


Fig. 1. A possible index codebook construction by the Huffman approach when $n = 4$ and $k = 2$. The number inside each node (square) is the probability used to facilitate the Huffman procedure. The sets beneath the leaf nodes represent the SAPs, i.e., $\{1, 2\}$ means that the first and the second of the four subcarriers are active. The bit sequences under the SAPs are the index domain messages that are encoded by the latter.

IM, where $\binom{n}{k}$ may not be a power of two. To design an index codebook that includes all SAPs, we construct a binary tree T_h with $\binom{n}{k}$ leaves, where each leaf represents an SAP. Let p_i denote the probability that is assigned to the i th leaf, where $i = 1, 2, \dots, \binom{n}{k}$ and $\sum_{i=1}^{\binom{n}{k}} p_i = 1$. An internal node of T_h is obtained by merging two nodes that carry the lowest probabilities among all existing nodes that have not been subjected to merging; see [8] for more details.

After the tree construction, a bit is assigned to each of the edges, such that the two edges that share the same parent node cannot have the same bit value. The bit sequence encoded by the i th SAP is obtained by reading the bits along the path from the root to the i th leaf; Fig. 1 shows a particular example.

The probability that the i th SAP is used is equal to the probability that the bit sequence along the edges that link the root of T_h and the i th leaf is observed when the encoder reads the information bits. For a maximally entropic information source², the probability that the i th SAP occurs is 0.5 raised to the power of the depth of the i th leaf.

Remark 1. The probability p_i is assigned to direct the Huffman coding procedure and should not be confused with the probability that the i th SAP occurs during transmissions.

B. Throughput Enhancement in the Index Domain

We now analyze the throughput gain in the index domain by applying Huffman coding compared to the OFDM-IM approach in [4]. We do so by calculating the encoding rate of the involved schemes in the index domain.

For the OFDM-IM approach in [4], because all SAPs encode $\lfloor \log_2 \binom{n}{k} \rfloor$ bits, the index-domain encoding rate is

$$r = \left\lfloor \log_2 \binom{n}{k} \right\rfloor. \quad (1)$$

Let d_i denote the depth of the i th leaf of T_h . Then, using Huffman coding, the probability of observing the length- d_i bit sequence encoded by the i th SAP is 0.5^{d_i} , and the encoding rate of the index codebook constructed by T_h is

$$r_h = \sum_{i=1}^{\binom{n}{k}} 0.5^{d_i} d_i. \quad (2)$$

²A maximally entropic information source emits independent bits, where each bit has a probability of 0.5 of being either zero or one.

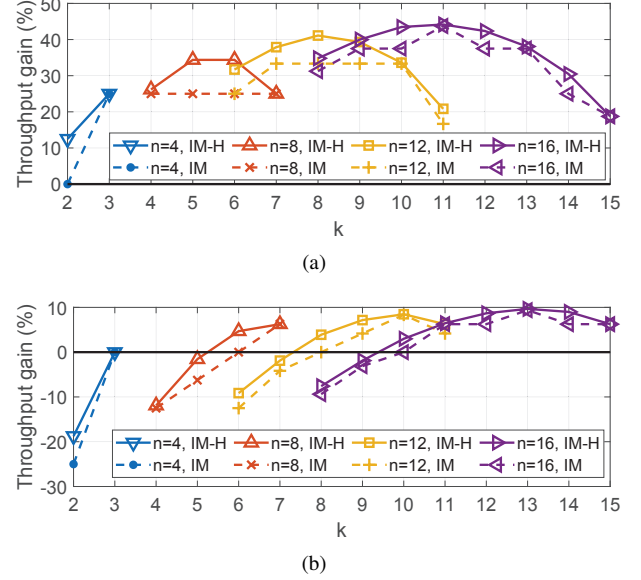


Fig. 2. Throughput gain of IM per n -subcarrier against conventional OFDM. In the legend, 'IM' means the scheme in [4] and 'IM-H' means Huffman-coding-aided IM with index-domain encoding rate r_h^* . (a) $M = 2$. (b) $M = 4$.

To see how to maximize r_h , we relax the integer constraint on d_i and take the derivative of r_h with respect to d_i , i.e.,

$$\frac{\partial r_h}{\partial d_i} = -\frac{\ln(2) d_i - 1}{2^{d_i}} \begin{cases} \geq 0, & \text{if } d_i \leq 1/\ln(2), \\ < 0, & \text{otherwise.} \end{cases} \quad (3)$$

Because d_i only takes nonnegative integer values, (3) indicates that we should keep d_i as small as possible for all i to maximize r_h . Then, to construct a binary tree with $\binom{n}{k}$ leaves of minimum depth, we can choose any D out of $2^{\lfloor \log_2 \binom{n}{k} \rfloor}$ leaves of a perfect binary tree and attach 2 child nodes to each of the chosen D leaves. This coincides with the design in [5], and is also equivalent to Huffman coding when $p_i = \binom{n}{k}^{-1}$ for all i ; see Appendix A for the equivalence between [5] and Huffman coding. Therefore, the maximum encoding rate of Huffman coding, denoted as r_h^* , is obtained when T_h has $2D$ leaves of depth $r+1$ and $\binom{n}{k} - 2D$ leaves of depth r , i.e.,

$$r_h^* = \left(\binom{n}{k} - 2D \right) \cdot 0.5^r \cdot r + 2D \cdot 0.5^{r+1} \cdot (r+1) \\ = r \cdot \left[\left(2^{r+1} - \binom{n}{k} \right) \cdot 0.5^r + \left(0.5^r \binom{n}{k} - 1 \right) \cdot \frac{r+1}{r} \right]. \quad (4)$$

Fig. 2 shows the throughput gain per n -subcarrier obtained by Huffman coded OFDM-IM and conventional OFDM-IM [4] compared to conventional OFDM. A 12.5% gain is achieved by using Huffman coding when $M = 2$, $n = 4$, and $k = 2$.

IV. MITIGATING SYNCHRONIZATION ERRORS

The variable-length nature of the Huffman coded IM introduces bit-level synchronization errors in the information bit sequence at the receiver output. For example, if a transmitted SAP that encodes r bits is detected incorrectly, such that the detector believes that another SAP encoding $r+1$ bits was sent, then one extra bit is erroneously inserted into the bit sequence at the detector output, which may significantly increase the

error events of the bits that follow the inserted bit. Coding techniques that correct synchronization errors exist [9], but significant rate loss may incur when applying them to correct synchronization errors in the context of this work.

Observe that the bit-level synchronization errors, when using a Huffman codebook with maximum encoding rate, can be eliminated if the detector knows whether r bits or $r + 1$ bits are encoded by a transmitted SAP. To inform the detector about this information *without any extra control signals*, we propose choosing one of two different constellations for any group of n subcarriers according to the number of bits encoded by the corresponding SAP. Specifically, let \mathcal{S} denote the set of all possible $\binom{n}{k}$ SAPs, and let $\mathcal{S}_{\text{short}}$ and $\mathcal{S}_{\text{long}}$ denote the sets of SAPs that encode r bits and $r + 1$ bits, respectively. Then, the transmission signal of a group of n subcarriers is taken from either $\mathcal{X}_{\text{short}} \triangleq \mathcal{S}_{\text{short}} \times \mathcal{M}_{\text{short}}^k$ or $\mathcal{X}_{\text{long}} \triangleq \mathcal{S}_{\text{long}} \times \mathcal{M}_{\text{long}}^k$, where $\mathcal{M}_{\text{short}}$ and $\mathcal{M}_{\text{long}}$ are the bandpass constellations when the SAP encodes r bits and $r + 1$ bits, respectively. At the receiver, maximum likelihood (ML) detection can be applied to perform joint detection on the transmitted SAP and the signals modulated on the active subcarriers, where only the signals from $\mathcal{X}_{\text{short}}$ and $\mathcal{X}_{\text{long}}$ are tested for their likelihoods.

For combinatorial OFDM-IM [4] with Huffman coding, the above constellation adaptation can be applied by setting $\mathcal{M}_{\text{short}}$ as M -ary phase-shift keying (PSK)³, and then rotating $\mathcal{M}_{\text{short}}$ by π/M to obtain $\mathcal{M}_{\text{long}}$. Fig. 3(a) illustrates the constellation adaptation for $M = 2$ and $M = 4$. For ease of reference, we abbreviate the proposed scheme as OFDM-IM-H2C, where ‘‘H2C’’ stands for Huffman coding with 2 Constellations.

Let P_{ins} be the probability that an SAP carrying r bits is decoded as an SAP carrying $r + 1$. Also, let \mathbf{x} and \mathbf{y} denote the transmission symbols on a group of n subcarriers and the corresponding received symbols, respectively. Assuming PSK signaling with unit power and an additive white Gaussian noise channel, where $\mathbf{y} = \mathbf{x} + \mathbf{z}$ and $\mathbf{z} \sim \mathcal{CN}(0, 2\sigma^2 \mathbf{I}_{n \times n})$, we have

$$\begin{aligned} P_{\text{ins}} &= \Pr\{\mathbf{x} \in \mathcal{X}_{\text{short}}\} \cdot \Pr\{\hat{\mathbf{x}} \in \mathcal{X}_{\text{long}} \text{ and } \|\mathbf{y} - \mathbf{x}\|^2 > \|\mathbf{y} - \hat{\mathbf{x}}\|^2\} \\ &= 0.5^r M^{-k} \cdot \sum_{\hat{\mathbf{x}} \in \mathcal{X}_{\text{long}}} \Pr\{\|\mathbf{y} - \mathbf{x}\|^2 > \|\mathbf{y} - \hat{\mathbf{x}}\|^2\} \\ &= 0.5^r M^{-k} \cdot \sum_{\hat{\mathbf{x}} \in \mathcal{X}_{\text{long}}} \Pr\{2\Re\{(\mathbf{x} - \hat{\mathbf{x}})^* \cdot \mathbf{z}\} < -\|\mathbf{x} - \hat{\mathbf{x}}\|^2\} \\ &= 0.5^r M^{-k} \cdot \sum_{\hat{\mathbf{x}} \in \mathcal{X}_{\text{long}}} \Pr\{\varepsilon < -0.5\|\mathbf{x} - \hat{\mathbf{x}}\|\} \quad (5a) \\ &\leq 0.5^{r-1} D \cdot Q\left((\sqrt{2}\sigma)^{-1}\right), \quad (5b) \end{aligned}$$

where $Q(\cdot)$ is the Q-function, $(\cdot)^*$ denotes conjugate transpose, $\varepsilon \sim \mathcal{N}(0, \sigma^2)$, (5a) follows because $\Re\{(\mathbf{x} - \hat{\mathbf{x}})^* \cdot \mathbf{z}\} \sim \mathcal{N}(0, \|\mathbf{x} - \hat{\mathbf{x}}\|^2 \sigma^2)$, and (5b) follows because the Euclidean distance between any $\mathbf{x} \in \mathcal{X}_{\text{short}}$ and any $\hat{\mathbf{x}} \in \mathcal{X}_{\text{long}}$ is at least two. A similar upper bound on the probability that an SAP carrying $r + 1$ bits is decoded incorrectly as an SAP carrying r , denoted as P_{del} , can be obtained as

$$P_{\text{del}} \leq 0.5^{r+1} \left[\binom{n}{k} - 2D \right] \cdot Q\left(\frac{1}{\sqrt{2}\sigma}\right). \quad (6)$$

³We skip the discussion for high-order modulations such as quadrature amplitude modulation, because comparing to conventional OFDM, OFDM-IM improves the throughput only when the modulation order is low.

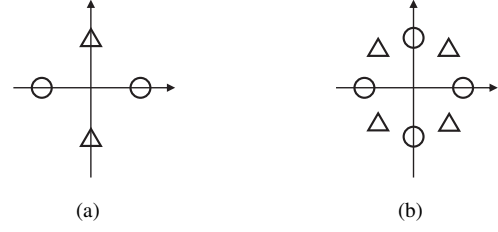


Fig. 3. PSK constellation adapted for the Huffman-coding-aided OFDM-IM to mitigate bit-level synchronization errors. Circles represent signals of $\mathcal{M}_{\text{short}}$, and triangles represent signals of $\mathcal{M}_{\text{long}}$. (a) $M = 2$. (b) $M = 4$.

Equation (5a) suggests that the proposed scheme mitigates bit-level synchronization errors by enlarging the Euclidean distance between a codeword in $\mathcal{X}_{\text{short}}$ and a codeword in $\mathcal{X}_{\text{long}}$.

V. SIMULATION RESULTS

The BERs of conventional OFDM, the OFDM-IM [4], the OFDM-IM with Huffman coding (OFDM-IM-H), the scheme in [5], and the proposed OFDM-IM-H2C are plotted against signal-to-noise ratio (SNR) in Fig. 4, where a Huffman codebook with maximum encoding rate is used when applicable, and ML detection is applied to all the schemes. For the scheme in [5], when $M = 2$, we always modulate the symbol ‘+1’ onto the last active subcarrier in a group of n subcarriers if the corresponding SAP encodes $r + 1$ index-domain bits. Also, for OFDM-IM-H2C, we count the bit errors in the index domain and the non-index domain separately by 1) splitting the input/output bit stream into two corresponding bit streams and 2) counting the bit errors in each domain as $\epsilon_s + \epsilon_e$, where

- 1) ϵ_s is the length difference between the encoded bit sequence and the decoded bit sequence in the corresponding domain, and,
- 2) ϵ_e is the number of bits in disagreement when comparing the encoded bit sequence and the decoded bit sequence by discarding the last ϵ_s bits of the former or the latter such that the two sequences have the same length.

Fig. 4(a) suggests that when applying Huffman coding to the OFDM-IM without treating the bit-level synchronization issue, a huge loss in error performance is observed (e.g., about 4-5 dB loss in terms of SNR at a BER of 10^{-3}). However, by applying the proposed constellation rotation method, the BER performance of the OFDM-IM-H2C becomes very close to that of conventional OFDM-IM in the high-SNR region, because the bit-synchronization errors have been reduced as shown in Fig. 5. This shows that the proposed scheme can improve the throughput without sacrificing much of the BER compared to OFDM-IM. Fig. 4(b) shows that the proposed OFDM-IM-H2C, when compared to all benchmarks, can have better throughput and lower BER in certain IM settings. The results in Fig. 2(b) and Fig. 4(c) suggest that $M = 2$ is of primary interest for (Huffman-coding-aided) OFDM-IM.

VI. CONCLUSION

We formally characterized the index codebooks based on Huffman coding that maximize the encoding rate in the index domain. When applying such an index codebook, to mitigate

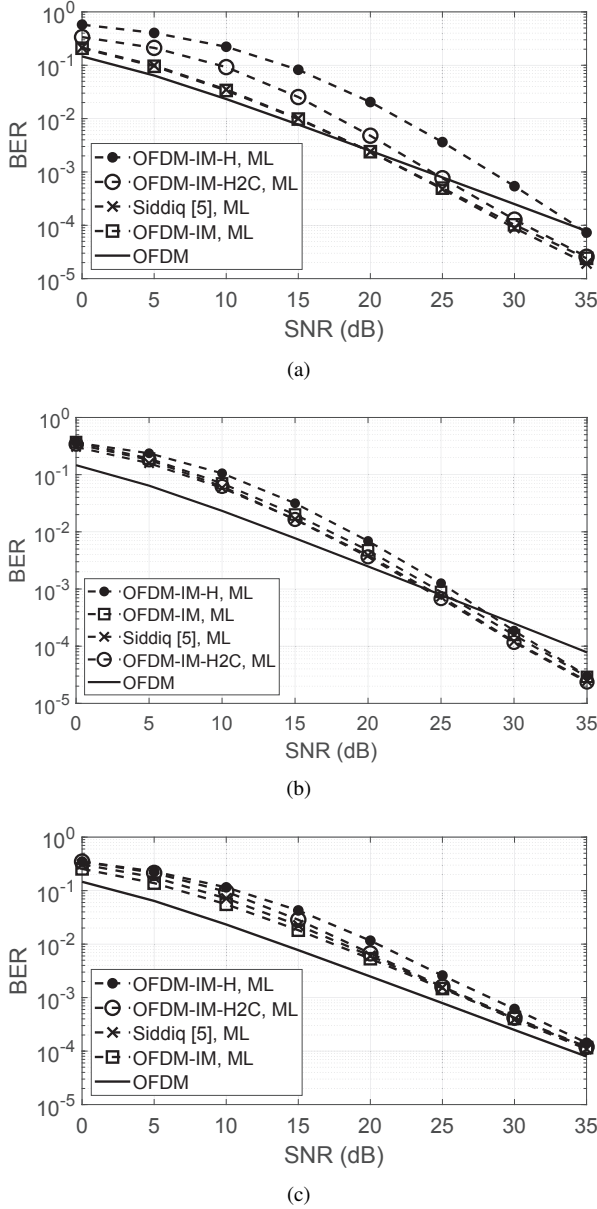


Fig. 4. BER performance comparison in 10-tap Rayleigh fading channel with 16 cyclic prefixes. (a) $M = 2$, $G = 32$, $n = 4$, and $k = 2$. (b) $M = 2$, $G = 16$, $n = 8$, and $k = 4$. (c) $M = 4$, $G = 21$, $n = 6$, and $k = 5$.

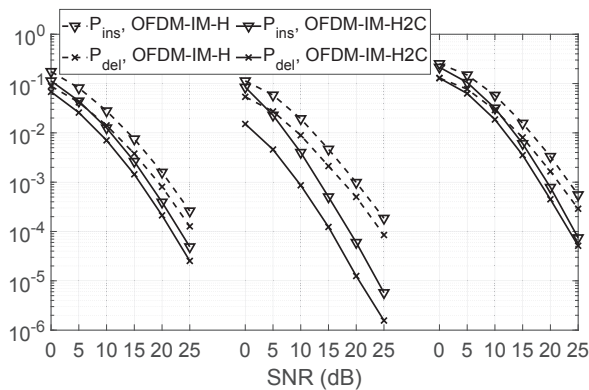


Fig. 5. Probabilities of bit-level synchronization errors with 10-tap Rayleigh fading channel and 16 cyclic prefixes. (a) $M = 2$, $G = 32$, $n = 4$, and $k = 2$. (b) $M = 2$, $G = 16$, $n = 8$, and $k = 4$. (c) $M = 4$, $G = 21$, $n = 6$, and $k = 5$.

bit-level synchronization errors, we proposed using two constellations for the bandpass signals to inform the receiver of the length of the index domain message without any signal overhead. Simulation results demonstrated that the proposed constellation adaptation scheme allows one to harvest the throughput gain brought by Huffman coding without degrading much of the error performance.

ACKNOWLEDGEMENTS

The authors would like to thank Dene A. Hedges for the useful discussions that lead to this work.

APPENDIX A

EQUIVALENCE BETWEEN THE HUFFMAN PROCEDURE AND THE EXTENSION OF THE PERFECT BINARY TREE

In [5], an index codebook is constructed by first selecting any D leaves from a perfect binary tree with $2^{\lceil \log_2(n) \rceil}$ leaves and then attach two child nodes to each of the selected leaves. This results in a tree with the following properties, i.e.,

- 1) the tree is a full binary tree,
- 2) there are $\binom{n}{k}$ leaves, and
- 3) a leaf has a depth of either $\lfloor \log_2 \binom{n}{k} \rfloor$ or $\lceil \log_2 \binom{n}{k} \rceil$.

The tree obtained by the Huffman coding approach, when assuming the $\binom{n}{k}$ source symbols happen with equal probability (i.e., $p_1 = p_2 = \dots = p_{\binom{n}{k}} = \binom{n}{k}^{-1}$), also satisfies the above three properties, because

- 1) every internal node has two child nodes so that the tree is full and binary, and the Huffman approach begins with $\binom{n}{k}$ leaves;
- 2) if the heights of subtrees that are subject to merging are either L or $L+1$, where L is a non-negative integer, then the Huffman approach will not merge two subtrees with heights $(L+1)$. This ensures that the depths of any two leaves cannot differ by two or more.

As a result, the index codebook construction method in [5] is equivalent to the Huffman coding procedure described in Section III when $p_1 = p_2 = \dots = p_{\binom{n}{k}} = \binom{n}{k}^{-1}$.

REFERENCES

- [1] D. Slepian, "Permutation modulation," *Proceedings of the IEEE*, vol. 53, no. 3, pp. 228–236, Mar 1965.
- [2] R. Abu-alhiga and H. Haas, "Subcarrier-index modulation OFDM," in *IEEE 20th International Symposium on Personal, Indoor and Mobile Radio Communications*, Sept 2009, pp. 177–181.
- [3] D. Tsonev, S. Sinanovic, and H. Haas, "Enhanced subcarrier index modulation (SIM) OFDM," in *IEEE GLOBECOM Workshops (GC Wkshps)*, Dec 2011, pp. 728–732.
- [4] E. Başar, U. Aygölü, E. Panayircı, and H. V. Poor, "Orthogonal frequency division multiplexing with index modulation," *IEEE Transactions on Signal Processing*, vol. 61, no. 22, pp. 5536–5549, Nov 2013.
- [5] A. I. Siddiq, "Low complexity OFDM-IM detector by encoding all possible subcarrier activation patterns," *IEEE Communications Letters*, vol. 20, no. 3, pp. 446–449, March 2016.
- [6] M. I. Kadir, H. Zhang, S. Chen, and L. Hanzo, "Entropy coding aided adaptive subcarrier-index modulated OFDM," *IEEE Access*, vol. 6, pp. 7739–7752, 2018.
- [7] Y. Yang, W. Cheng, W. Zhang, and H. Zhang, "Mode modulation for wireless communications with a twist," *IEEE Transactions on Vehicular Technology*, vol. 67, no. 11, pp. 10704–10714, Nov 2018.
- [8] D. A. Huffman, "A method for the construction of minimum-redundancy codes," *Proceedings of the IRE*, vol. 40, no. 9, pp. 1098–1101, Sept 1952.
- [9] M. Mitzenmacher, "A survey of results for deletion channels and related synchronization channels," *Probability Surveys*, vol. 6, pp. 1–33, 2009.