



## **Multilevel Coding with Flexible Probabilistic Shaping for Rate-Adaptive and Low-Power Optical Communications**

Downloaded from: <https://research.chalmers.se>, 2026-04-03 18:04 UTC

Citation for the original published paper (version of record):

Yoshida, T., Karlsson, M., Agrell, E. (2020). Multilevel Coding with Flexible Probabilistic Shaping for Rate-Adaptive and Low-Power Optical Communications. 2020 Optical Fiber Communications Conference and Exhibition, OFC 2020 - Proceedings. <http://dx.doi.org/10.1364/OFC.2020.M3J.7>

N.B. When citing this work, cite the original published paper.

# Multilevel Coding with Flexible Probabilistic Shaping for Rate-Adaptive and Low-Power Optical Communications

Tsuyoshi Yoshida<sup>1,2</sup>, Magnus Karlsson<sup>3</sup>, and Erik Agrell<sup>3</sup>

<sup>1</sup> Information Technology R&D Center, Mitsubishi Electric Corporation, Kamakura, 247-8501 Japan

<sup>2</sup> Graduate School of Engineering, Osaka University, Suita, 505-0871 Japan

<sup>3</sup> Fiber-Optic Communications Research Center (FORCE), Chalmers University of Technology, SE-41296 Gothenburg, Sweden  
Yoshida.Tsuyoshi@ah.MitsubishiElectric.co.jp, yoshida.tsuyoshi@opt.comm.eng.osaka-u.ac.jp

**Abstract:** A novel multilevel coded modulation scheme with probabilistic shaping is presented. It can reduce the power consumption up to 9 times compared with uniform signaling in the regime of typical hard-decision FEC thresholds. © 2020 The Author(s)  
**OCIS codes:** (060.2330) Fiber Optics Communication; (060.4080) Modulation

## 1. Introduction

The optical communications community is moving forward to wide deployment of 400 Gb/s per optical carrier based on standardizations such as 400ZR and openROADM. For such systems with high baud-rate signaling, quadrature amplitude modulation (QAM) and efficient soft forward error correction (FEC) are key techniques. Beyond 400 Gb/s, probabilistic shaping (PS) could be a new important technique for reach extension and rate adaptability [1,2]. PS has mainly been studied together with bit-interleaved coded modulation (BICM) [3]. The BICM-PS scheme realizes capacity-approaching performance, but causes an FEC throughput increase which leads to large power consumption [4,5]. To reduce the power consumption, joint source–channel coding [6], shallow shaping having granular base constellations [7], and multilevel coding (MLC) [8–10] have been investigated. While MLC with multi-stage decoding (MSD) has received attention [10,11], MLC-MSD with PS has not been studied in depth yet. A major drawback with MLC is that in order to realize rate flexibility, the FEC code rate of each bit tributary needs to be adapted, which makes the system complex. However, now PS can provide such rate flexibility. In this work we investigate the combination of almost fixed MLC and flexible PS, enabling rate adaptation and lower power consumption than BICM.

## 2. Principle of MLC-PS

The proposed MLC-PS scheme is explained in this section. Since the power consumption with soft FEC is significantly larger than with hard FEC, we protect only one bit tributary by soft FEC, to reduce the power consumption relative to, e.g., [10]. We combine two symbol mapping rules as described in [10], which improves the performance. Fig. 1 shows the block diagram of the proposed MLC-PS scheme with an exemplified maximum constellation size of 256-QAM. At the transmitter side (a), first, PS encoding generates amplitude bits  $B_a$  from source bits  $S_a$  based on the symbol mapping  $\mathbf{B}^{(k)} \rightarrow X$ , which is symmetric in positive and negative amplitudes (see below). Next,  $B_s$ , from source bits  $S_s$ , and  $B_a$  are multiplexed and encoded by a single systematic FEC encoder into  $\mathbf{C} = C_1^{(1)} C_2^{(1)} C_3^{(1)} C_4^{(1)} C_1^{(2)} C_2^{(2)} C_3^{(2)} C_4^{(2)}$ , where the bits  $C_1^{(k)}$  are not sent because of fixed to zeros. The bit  $C_4^{(k)}$  is XOR'ed with  $C_1^{(k)}$  to give  $B_4^{(k)}$ . Then,  $B_4^{(1)}$  and  $B_4^{(2)}$  ( $= C_4^{(2)}$ ) are multiplexed and encoded by a systematic FEC (for soft decoding). The generated parity is assigned to  $B_1^{(2)}$ . Note that in PS-64- and 16-QAM zeros are padded in non-used bit tributaries in  $B_a$ . Then, e.g.,  $C_2^{(k)}$ ,  $C_3^{(k)}$   $B_2^{(k)}$ , and  $B_3^{(k)}$  are all zero in PS-16-QAM. At the receiver side (b), first,  $\hat{B}_4^{(1)}$  and  $\hat{B}_4^{(2)}$  are decoded by soft FEC from a *posteriori* L-values  $L_4^{(1)}$ ,  $L_4^{(2)}$ , and  $L_1^{(2)}$ . The other bits are preliminarily decoded to  $\hat{B}_1^{(1)}$ ,  $\hat{B}_2^{(k)}$ , and  $\hat{B}_3^{(k)}$  ( $k = 1, 2$ ) for the cases of  $\hat{B}_4^{(k)} = b$  ( $b = 0, 1$ ). After latency adjustment,  $\hat{B}_1^{(1)}$ ,  $\hat{B}_2^{(k)}$ , and  $\hat{B}_3^{(k)}$  yield  $\hat{C}_1^{(1)}$ ,  $\hat{C}_2^{(k)}$ , and  $\hat{C}_3^{(k)}$  by hard decisions based on  $\hat{B}_4^{(k)}$ . Then the XOR operation at the transmitter is

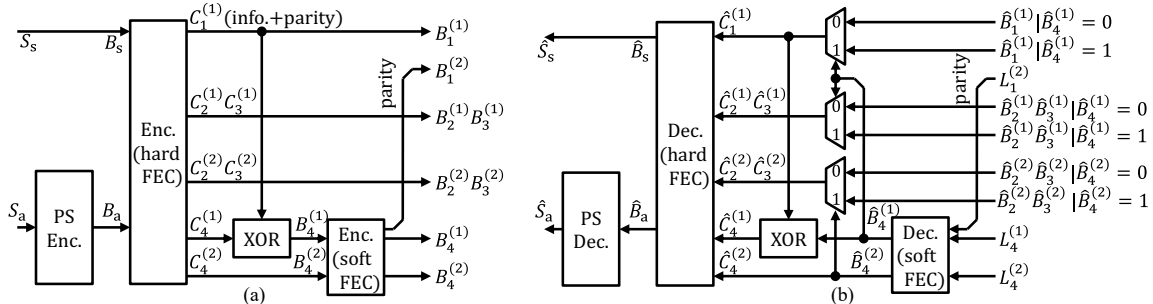


Fig. 1. Functional block diagrams of (a) transmitter and (b) receiver side MLC-PS processing for 256-QAM ( $m_n = 8$ ).

reverted to  $\hat{C}_4^{(1)}$ , and  $\hat{B}_s (= \hat{S}_s)$  and  $\hat{B}_a$  are decoded from  $\hat{C}$  by hard FEC with a long codeword, where zeros are padded in  $\hat{C}_1^{(2)}$  inside the soft FEC decoding. Finally PS decoding is performed for  $\hat{B}_a$  to recover  $\hat{S}_a$ .

Fig. 2 shows the frame structure of our MLC-PS, using PS-256-QAM as an example. Real-valued symbols  $X$  are generated from groups of four bits according to Tab. 1, and QAM symbols are generated by combining two real-valued symbols. The bits  $B_1^{(k)}-B_3^{(k)}$  are binary reflected Gray coded,  $B_4^{(1)}$  is used for the set partitioning of amplitude levels (16 levels to two sets of 8 levels), and  $B_4^{(2)}$  is used for the set partitioning of absolute amplitude levels (8 levels to two sets of 4 levels). In the case  $k = 2$ , complete set partitioning is not required because the sign bit is not used after soft FEC decoding. This helps to reduce the performance degradation by the set partitioning. For simplicity we put zeros in non-used bit tributaries, and the hard FEC parity is not protected by the soft FEC. Then 64-QAM and 16-QAM use the amplitudes  $-7$  to  $7$  and  $-3$  to  $3$ , resp. As with probabilistic amplitude shaping (PAS) [1], all parity bits of soft and hard FECs are placed on the sign bits  $B_1^{(1)}$  and  $B_1^{(2)}$ , resp. Note that the soft FEC parity is actually placed on  $B_1^{(2)}$  of the next soft FEC frame to disperse burst errors. The information rate with the proposed MLC-PS scheme is

$$R_{\text{mlc-ps}} = (m - 2)R_{\text{pc}} + 2(1 - r_{\text{sc}}) - m_0(1 - R_{\text{hc}}), \quad (1)$$

where  $m$  and  $m_0$  are the effective and the maximum numbers of bit slots per two-dimensional symbol, resp. For example, in the case of PS-64-QAM with a PS-256-QAM MLC frame,  $m = 6$  and  $m_0 = 8$ . The rate  $R_{\text{pc}}$  refers to the PS code rate (ratio of the numbers of the input and output bits). The  $r_{\text{sc}}$  is soft FEC redundancy of  $1/R_{\text{sc}} - 1$ . The  $R_{\text{sc}}$  and  $R_{\text{hc}}$  are soft and hard FEC code rates, resp. For comparison, the information rate of BICM-PS is

$$R_{\text{bicm-ps}} = (m - 2)R_{\text{pc}} + 2 - m(1 - R_{\text{c}}), \quad (2)$$

where  $R_{\text{c}}$  is total FEC code rate including soft and hard FECs. Eqs. (1) and (2) will be used in Sec. 3 to choose  $R_{\text{pc}}$ .

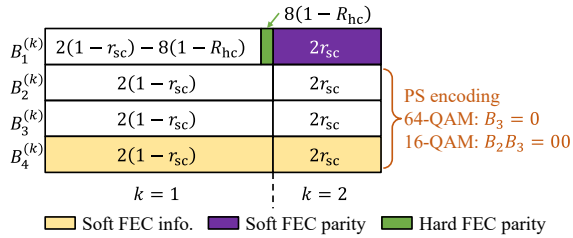


Fig. 2. Frame structure of the proposed MLC-PS with 256-QAM. Expressions inside the squares denote the normalized number of bit slots per QAM symbol (8-bit space of the 256-QAM symbol).

### 3. Rate adaptation, performance, and power consumption

We set target information rates of 5.6, 4.8, 4.0, and 3.2 bit per channel use (bpcu). These can be realized by BICM with uniform 128, 64, 32, and 16-QAM and an FEC code rate  $R_{\text{c}} = 4/5$ . In Tab. 2, the examined cases with our proposed MLC-PS and BICM-PS (PAS) are summarized. Here we assumed the use of hard FEC having a code rate  $R_{\text{hc}} = 0.99$ . The soft FEC code rate  $R_{\text{sc}}$  was chosen from  $2/3$ ,  $3/4$ ,  $4/5$ , or  $5/6$ . To achieve the target information rate, we adjust the PS code rate  $R_{\text{pc}}$  based on (1) and (2), where each PS frame encodes 128 PS-256-QAM or PS-64-QAM symbols, or 256 PS-16-QAM symbols. For the PS encoding, we employed hierarchical distribution matching to form the probability mass function of the symbols [5,12]. The shaping gap from the ideal Maxwell-Boltzmann distribution is less than 0.3 dB. The relative values of the decoder input rate and the power consumption,  $T_{\text{sc}}$  and  $P_{\text{sc}}$ , in the soft FEC decoding at the same system throughput (payload rate) are also described in Tab. 2. These were normalized by the values in the case of uniform QAM with  $R_{\text{sc}} = 4/5$ . The baseline power consumption is proportional to circuit size, clock frequency, and number of instances. Here we assume that the FEC codeword length and the number of instances

Tab. 2. FEC/PS combinations: proposed MLC with PS and BICM with PS for rate adaptation.

	5.6 bpcu				4.8 bpcu				4.0 bpcu				3.2 bpcu				3.2 bpcu			
$R_{\text{sc}}$	2/3	3/4	4/5	5/6	2/3	3/4	4/5	5/6	2/3	3/4	4/5	5/6	2/3	3/4	4/5	5/6	2/3	3/4	4/5	5/6
	MLC PS-256-QAM				MLC PS-64-QAM				MLC PS-64-QAM				MLC PS-64-QAM				MLC PS-16-QAM			
$R_{\text{pc}}$	0.78	0.73	0.70	0.68	0.97	0.89	0.85	0.82	0.77	0.69	0.65	0.62	0.57	0.49	0.45	0.42	-	0.97	0.89	0.84
$T_{\text{sc}}$	0.43	0.38	0.36	0.34	0.50	0.44	0.42	0.40	0.60	0.53	0.50	0.48	0.75	0.67	0.62	0.60	-	0.67	0.62	0.60
$P_{\text{sc}}$	0.86	0.51	0.36	0.34	1.00	0.59	0.42	0.32	1.20	0.71	0.50	0.38	1.50	0.89	0.62	0.48	-	0.89	0.62	0.48
	BICM PS-256-QAM				BICM PS-64-QAM				BICM PS-64-QAM				BICM PS-64-QAM				BICM PS-16-QAM			
$R_{\text{pc}}$	-	-	0.88	0.83	-	-	-	0.96	-	0.89	0.81	0.76	-	0.69	0.61	0.56	-	-	-	0.95
$T_{\text{sc}}$	-	-	1.14	1.14	-	-	-	1.00	-	1.20	1.20	1.20	-	1.50	1.50	1.50	-	-	-	1.00
$P_{\text{sc}}$	-	-	1.14	0.91	-	-	-	0.80	-	1.60	1.20	0.96	-	2.00	1.50	1.20	-	-	-	0.80

Tab. 1. Symbol mapping  $\mathbf{B}^{(k)} \rightarrow X$  for MLC-PS, where  $\mathbf{B}^{(k)} = B_1^{(k)}B_2^{(k)}B_3^{(k)}B_4^{(k)}$ .

$X$	$\mathbf{B}^{(1)}$	$\mathbf{B}^{(2)}$	$X$	$\mathbf{B}^{(1)}$	$\mathbf{B}^{(2)}$
15	0101	0101	-1	1001	1000
13	0100	0100	-3	1000	1001
11	0111	0111	-5	1011	1010
9	0110	0110	-7	1010	1011
7	0011	0011	-9	1111	1110
5	0010	0010	-11	1110	1111
3	0001	0001	-13	1101	1100
1	0000	0000	-15	1100	1101

are respectively the same among the examined cases. Then, as a rough estimation, the power consumption is proportional to the number of soft FEC parity bits and the decoder input rate. As a result, MLC-PS shows around 2 times lower power consumption than BICM-PS at the same information rate and  $R_{sc}$ .

Fig. 3 shows the simulated bit error rate (BER) before the hard FEC decoding over the Gaussian channel. As the soft FEC, we employed DVB-S2 low-density parity check code having an  $R_{sc}$  of 2/3, 3/4, 4/5, or 5/6 and a codeword length of 64800 bits. The maximum number of decoding iterations was set to 20. When  $R_{sc}$  is smaller in the case of MLC-PS, the waterfall performance is better if PS code rate is sufficiently small, e.g.,  $R_{pc} < 0.95$ , and the error floor performance is worse due to the shallower PS. Especially in the case of MLC-PS with  $R_{sc} = 2/3$ , the error floor due to the MSD is  $4 \cdot 10^{-4}$ , which leads to rather large complexity in a hard FEC with  $R_{hc}$  of 0.99 to clean up to error free ( $BER < 10^{-15}$ ). Fig. 4 shows the relative power consumption, where we assumed that the soft decoding power consumption is proportional to the average decoding iterations. We also considered the relative power consumption in Tab. 2 here. From Figs. 3 and 4, though the performance of MLC-PS is slightly worse than BICM-PS, the power consumption is typically 2–3 times smaller at hard-decision FEC thresholds around  $10^{-4}$ . Compared with uniform signaling, the power consumption is about 9 times less for 5.6 bpcu at a signal-to-noise ratio (SNR) of 19 dB.

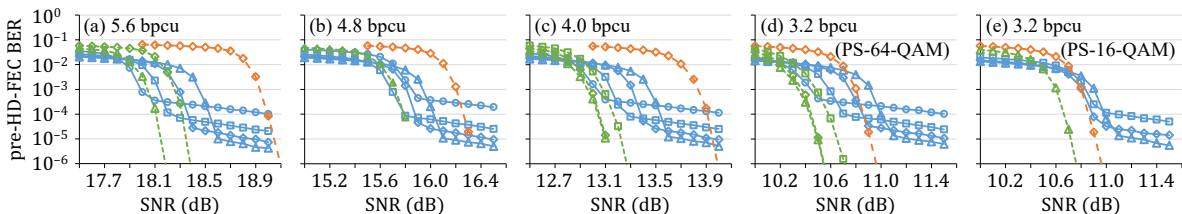


Fig. 3. Simulated BER before hard decoding for each information rate (bpcu) and base constellation cases: solid blue line: MLC-PS, dotted green line: BICM-PS, dashed orange line: BICM with uniform signaling, circle:  $R_{sc}=2/3$ , square:  $R_{sc}=3/4$ , diamond:  $R_{sc}=4/5$ , and triangle:  $R_{sc}=5/6$ .

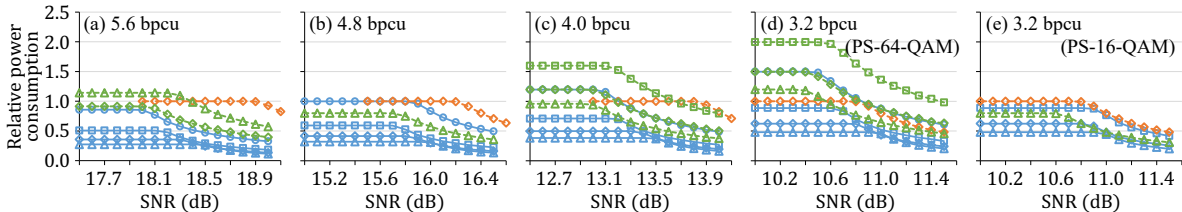


Fig. 4. Relative power consumption of soft FEC decoding. The subfigure names and the legends are the same as in Fig. 3.

#### 4. Conclusions

We proposed the combination of MLC and flexible PS, and compared the performance and the power consumption among relevant coded modulation schemes. MLC-PS realizes less power consumption by typically 2–3 times than the state-of-the-art BICM-PS and by at maximum 9 times than BICM with uniform signaling because of the lower soft FEC throughput and fewer decoding iterations.

#### Acknowledgments

This work was partly supported by “The research and development of innovative optical network technology as a new social infrastructure” of the Ministry of Internal Affairs and Communications, “Massively Parallel and Sliced Optical Network,” the Commissioned Research of National Institute of Information and Communications Technology (NICT), Japan, and the Swedish Research Council under grant 2017-03702. We also thank Assoc. Prof. Koji Igarashi and Prof. Kyo Inoue of Osaka University for assistance in the research.

#### References

- [1] G. Böcherer *et al.*, “Bandwidth efficient and rate-matched ...,” *IEEE Trans. Commun.*, **63**(12), 4651–4665, 2015.
- [2] F. Buchali *et al.*, “Rate adaptation and reach increase by ...,” *J. Lightw. Technol.*, **34**(7), 1599–1609, 2016.
- [3] G. Caire *et al.*, “Bit-interleaved coded modulation,” *IEEE Trans. Inf. Theory*, **44**(3), 873–884, 1992.
- [4] T. Yoshida *et al.*, “Technologies toward implementation of probabilistic ...,” in *Proc. ECOC*, 2018, Paper Th.1.H.1.
- [5] T. Yoshida *et al.*, “Hierarchical distribution matching for ...,” *J. Lightw. Technol.*, **37**(6), 1579–1589, 2019.
- [6] T. Yoshida *et al.*, “Joint source–channel coding via compressed probabilistic shaping ...,” in *Proc. OFC*, 2019, Paper M4B.6.
- [7] T. Yoshida and N. Suzuki, “Flexible and low-power probabilistic ...,” in *Proc. SPPCom*, 2019, Paper SpT3E.2.
- [8] L. Beygi *et al.*, “Adaptive coded modulation for nonlinear fiber-optic channels,” in *Proc. Globecom Workshops*, 2012.
- [9] O. İşcan *et al.*, “Probabilistically shaped multi-level coding with polar codes for fading ...,” in *Proc. Globecom Workshops*, 2018.
- [10] K. Sugitani *et al.*, “Partial multilevel coding with probabilistic shaping for ...,” in *Proc. OECC/PSC*, 2019, Paper TuB1-5.
- [11] F. Frey *et al.*, “Coded modulation using a 512-ary Hurwitz-integer constellation,” in *Proc. ECOC*, 2019, Paper W.2.D.2.
- [12] T. Yoshida *et al.*, “FPGA implementation of distribution matching and ...,” in *Proc. ECOC*, 2019, Paper M.2.D.2.