



# CHALMERS

## Chalmers Publication Library

### **Post-FEC BER prediction accuracy for probabilistically shaped signaling in fiber-optic communications**

This document has been downloaded from Chalmers Publication Library (CPL). It is the author's version of a work that was accepted for publication in:

**European Conference on Optical Communication (ECOC)**

Citation for the published paper:

Yoshida, T. ; Karlsson, M. ; Agrell, E. (2017) "Post-FEC BER prediction accuracy for probabilistically shaped signaling in fiber-optic communications". European Conference on Optical Communication (ECOC)

Downloaded from: <http://publications.lib.chalmers.se/publication/254769>

Notice: Changes introduced as a result of publishing processes such as copy-editing and formatting may not be reflected in this document. For a definitive version of this work, please refer to the published source. Please note that access to the published version might require a subscription.

Chalmers Publication Library (CPL) offers the possibility of retrieving research publications produced at Chalmers University of Technology. It covers all types of publications: articles, dissertations, licentiate theses, masters theses, conference papers, reports etc. Since 2006 it is the official tool for Chalmers official publication statistics. To ensure that Chalmers research results are disseminated as widely as possible, an Open Access Policy has been adopted. The CPL service is administrated and maintained by Chalmers Library.

(article starts on next page)

# Post-FEC BER Prediction Accuracy for Probabilistically Shaped Signaling in Fiber-Optic Communications

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

<sup>(1)</sup> Chalmers University of Technology, SE412-96 Gothenburg, Sweden, [tsuyoshi@chalmers.se](mailto:tsuyoshi@chalmers.se)

<sup>(2)</sup> Mitsubishi Electric Corp., 5-1-1 Ofuna, Kamakura, Kanagawa, 247-8501 Japan, [Yoshida.Tsuyoshi@ah.MitsubishiElectric.co.jp](mailto:Yoshida.Tsuyoshi@ah.MitsubishiElectric.co.jp)

**Abstract** We analyze performance metrics to predict the post-FEC BER for probabilistically-shaped signaling. Numerical simulations over a fiber-optic channel show that the metric based on symmetrized and mixed LLRs predicts  $>10^2$  times more accurate post-FEC BER than GMI.

## Introduction

Probabilistic shaping has been investigated for optical communications to reduce the required signal-to-noise ratio (SNR) of uniformly distributed quadrature amplitude modulation (QAM) signals<sup>1</sup>. We usually use soft-decision (SD) forward error correction (FEC) due to the significant coding gains of more than 10 dB<sup>2</sup>. To quantify the achievable rate of a system with uniform signaling, the generalized mutual information (GMI) is helpful in bit-interleaved coded modulation. The bit error rate (BER) after SD-FEC decoding is predictable by the normalized GMI for uniform signals<sup>3,4</sup>. According to our recent study<sup>5</sup> of performance metrics for nonuniform (probabilistically shaped) signals on the Gaussian channel, the original normalized GMI fails to predict the post-FEC BER, and instead the asymmetric information (ASI) based on the concept of symmetrized and mixed log-likelihood ratios (LLRs)<sup>3,4</sup> succeeds better. However, this metric has not been examined on a realistic fiber-optic channel including digital signal processing or analog distortions such as laser phase noise. In this paper, we summarize the performance metrics for nonuniform signals, and show that ASI is the most accurate to predict the post-FEC BER also over the nonlinear fiber-optic channel by numerical simulations.

## System model and performance metrics

The system model for nonuniform signalling to be considered is shown in Figure 1. A uniform source signal is processed in a distribution matcher and converted into a nonuniform signal  $A$ , and encoded into  $B$  by a binary FEC with a bit interleaver. It is demultiplexed into bit channels  $\mathbf{B} = \{B_1, B_2, \dots, B_m\}$ , which via a mapper generates a symbol  $X$ . A received symbol  $Y$  from a channel is demapped to a bit-wise LLR

$$L_i = \ln \frac{\sum_{\mathbf{b}:b_i=0} P_{\mathbf{B}}(\mathbf{b}) p_{Y|\mathbf{B}}(y|\mathbf{b})}{\sum_{\mathbf{b}:b_i=1} P_{\mathbf{B}}(\mathbf{b}) p_{Y|\mathbf{B}}(y|\mathbf{b})}, \quad (1)$$

where the channel  $p_{Y|\mathbf{B}}(y|\mathbf{b})$  is typically approximated by a memoryless Gaussian

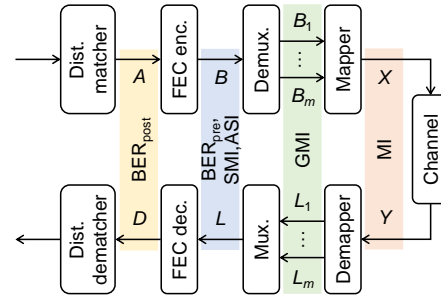


Fig. 1: System model.

channel in the demapper<sup>1</sup>. The LLRs are multiplexed to  $L$ , and decoded by a FEC decoder. The FEC encoder for nonuniform signals must be systematic, not to change the nonuniformity. For uniform signals, the distribution matcher and the dematcher are not required.

In real systems, the pre-FEC BER,  $BER_{pre}$ , is typically estimated by observing the number of flipped bits in the FEC decoder and converted into the Q-factor  $Q_{pre} = \sqrt{2} \text{erfc}^{-1}(BER_{pre})$  to characterize the system margin. For a signal having uniform and independent distributions over bit channels, the normalized GMI<sup>3,4</sup>  $I_g = \sum_{i=1}^m I(B_i; L_i)/m$  is regarded as a precise metric to predict the post-FEC BER,  $BER_{post}$ . It can also be described using the single-bit variables  $B, L$  as<sup>5</sup>  $I(B; L)$ , which we call *single-bit MI*. For nonuniform signals, to make the definition range  $[0,1]$ , we normalize the metric as

$$I_g = \sum_{i=1}^m I(B_i; L_i) / \sum_{i=1}^m H(B_i), \quad (2)$$

where  $H(\psi)$  denotes the entropy of a discrete variable  $\psi$ . This generalized definition is consistent with the regular GMI definition mentioned above for uniform signaling, because  $\sum_{i=1}^m H(B_i) = m$  in this case.

We will also consider the *asymmetric LLR*, which is  $L$  flipped by the transmitted bit  $B$  as<sup>5</sup>

$$L_a = (-1)^B L. \quad (3)$$

The concept has been referred to as a “symmetrized” and “mixed” L-values<sup>3</sup>. Here, the *asymmetric information (ASI)*<sup>5</sup>

$$I_a = 1 - h(L_a | |L_a|) = 1 - h(L_a) + h(|L_a|) \quad (4)$$

provides a quantitative value to measure the asymmetry of  $L_a$ , where  $h(\xi)$  denotes the differential entropy of a continuous variable  $\xi$ . The metrics  $I_g$  and  $I_a$  are equivalent and defined in  $[0,1]$  for uniformly, independently, and symmetrically distributed signals.

In practical receiver implementations, the LLRs are always quantized with a finite resolution, and can therefore be treated as discrete values. We adopt the same discretization in our analysis, which greatly facilitates the estimation of differential entropy and mutual information. Then  $I(B_i; L_i)$  or  $I(B; L)$  is calculated as

$$\hat{I}(B; L) = \sum_{\beta, \lambda} P_{B,L}(\beta, \lambda) \log_2 \frac{P_{B,L}(\beta, \lambda)}{P_B(\beta)P_L(\lambda)}, \quad (5)$$

and  $I_a$  as

$$\hat{I}_a = 1 - \sum_{\lambda} P_{L_a}(\lambda) \log_2 P_{L_a}(\lambda) + \sum_{\lambda} P_{|L_a|}(|\lambda|) \log_2 P_{|L_a|}(|\lambda|), \quad (6)$$

where  $\beta$  and  $\lambda$  are the ensembles of transmitted bits and the quantized LLRs. All probabilities in (5) and (6) are estimated from discrete histograms with finite range, which are based on the ensembles from either simulation or experiment.

### Simulation conditions

We have simulated the relations between pre-FEC metrics and post-FEC BERs by utilizing DVB-S2 low-density parity check (LDPC) codes<sup>6</sup> having a code length of 64800 and code rates of 2/3, 3/4, 5/6, and 9/10. The number of iterations in the decoding was 20. The parameters of the probabilistically-shaped 64-QAM signals evaluated are listed in Table 1. The shaping was done for 8-ary pulse amplitude modulation (8-PAM) symbols per dimension by probabilistically amplitude shaping<sup>7</sup> with the constant composition distribution matching<sup>8</sup>. The output block length  $N_s$  of the shaping was set to 1024, and the target probability mass functions were the same as in Tab.1(a)–(c) of<sup>9</sup>.  $N_{B_i}$  shows the number of input bits of bit tributary  $i$  to the distribution matcher per  $N_s$ . A uniform 64-QAM system was also simulated for comparison. The channel was either the Gaussian channel or a polarization-multiplexed fiber-optic channel including digital signal processing for a coherent detection and analog distortions at the optical transmitter and receiver, e.g. the linewidths and the carrier frequency offset of the lasers were 100 kHz in each and 500 MHz, resp. The resolution and the effective number of bits in digital-to-analog and analog-to-digital converters were 8 bits and 6 bits, resp. The adaptive equalization ( $T/2$ -

Tab. 1: Tested signal parameters for 64-QAM.

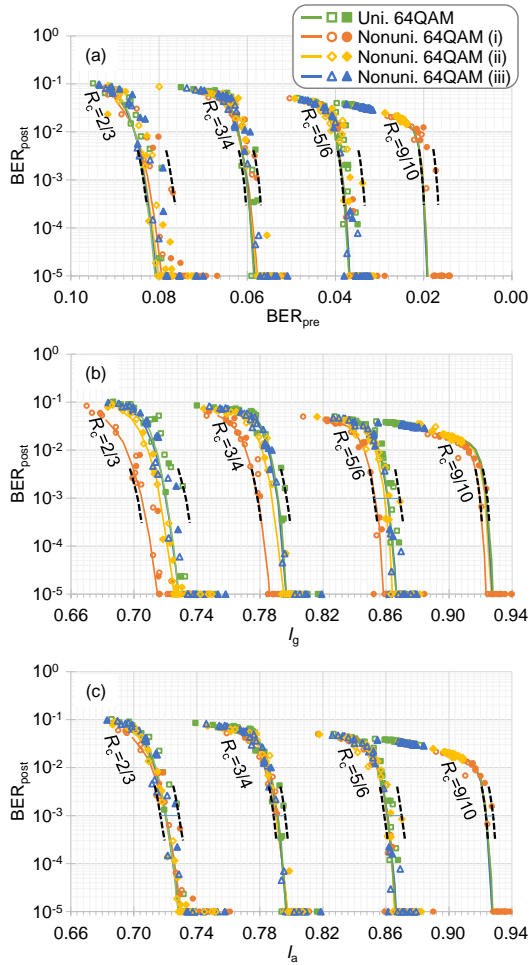
Condition	Uni.	Nonuni. (i)	Nonuni. (ii)	Nonuni. (iii)
target $P_{X 1}$ (1)	0.250	0.414	0.326	0.274
target $P_{X 1}$ (3)	0.250	0.316	0.290	0.264
target $P_{X 1}$ (5)	0.250	0.186	0.226	0.244
target $P_{X 1}$ (7)	0.250	0.084	0.158	0.218
$(N_{B_2}, N_{B_3})/N_s$	2	1.788	1.934	1.979
$\sum_i H(B_i)$	3	2.841	2.960	2.996
$H(B)$	1	0.9832	0.9968	0.9997

spaced, 21-tap) and the carrier recovery at the receiver side were performed by fully pilot-aided processings with 4% redundancy. The transmitted signal spectrum was shaped by a root raised cosine low-pass filter having a roll-off ratio of 1%. The symbol rate was 32 Gsymbol/s. Seven-channel wavelength division multiplexing was assumed with a 32.32 GHz spacing. Standard single-mode fibers were used, whose span length, number of spans, chromatic dispersion, loss, and nonlinear coefficient were respectively 100 km, 5 spans, 17 ps/nm, 0.2 dB/km, and  $1.2 \text{ W}^{-1}\text{km}^{-1}$ . An amplified spontaneous emission noise was loaded in each optical amplifier having a noise figure of 5 dB. The Manakov equation-based split-step Fourier method was used for the simulation of nonlinear propagation with adaptive steps. The launch power was varied from -13 to +7 dBm/ch.

### Simulation results

The metrics ( $\text{BER}_{\text{pre}}, I_g, I_a$ ) vs.  $\text{BER}_{\text{post}}$  are shown in Fig. 2. The solid lines show the performances over the Gaussian channel (averaging over at least 500 LDPC codewords) and the markers show the ones over the fiber-optic channel. The maximum peak-to-peak metric differences at  $\text{BER}_{\text{post}}$  of  $10^{-3}$  are 0.0024, 0.016, and 0.0014 for  $\text{BER}_{\text{pre}}, I_g$ , and  $I_a$ , resp., over the Gaussian channel<sup>5</sup>. Another extension of the normalized GMI for nonuniform signal is<sup>10</sup>  $(\sum_{i=1}^m I(B_i; L_i) + m - 1 - H(|X|))/m$ . We believe that “+1” in<sup>10</sup> is a typo and should be “-1” as above. Although this takes similar values as  $I_a$  under the operating conditions of typical probabilistic amplitude shaping schemes<sup>7</sup>, the definition range depends on  $H(|X|)$ , which is different from the others metrics.

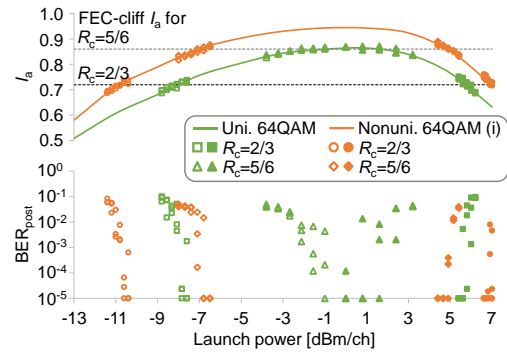
In the case of the fiber-optic channel, the pre-FEC metrics deviate between different simulations due to different laser phase noise realizations and the convergence of the adaptive equalization. The deviation is observed as well in Fig. 3, showing  $I_a$  (lines and markers are typical and single simulation performances, resp.) and  $\text{BER}_{\text{post}}$  (markers) vs. launch power for uniform 64-QAM and case (i) of nonuniform 64-QAM signaling. Thus we plotted single-simulation result (each single-simulation has 5 LDPC



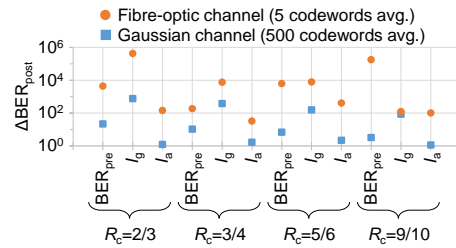
**Fig. 2:** Metric (a) Pre-FEC BER  $BER_{pre}$ , (b) normalized GMI  $I_g$ , or (c) ASI  $I_a$  vs. post-FEC BER on the Gaussian (solid lines) or the fibre-optic channel (markers).

codewords) without averaging them in Figs. 2 and 3. When  $BER_{post}$  is  $< 10^{-5}$ , the markers are shown at  $10^{-5}$ . The filled and the open markers correspond to linear and nonlinear conditions, resp. In Fig. 2, the markers (the fiber-optic channel) are basically correlated with the solid lines (the Gaussian channel). The deviation from the solid lines are due to the lack of averaging or the burst-error in the fiber-optic channel simulations (the deviation range around  $BER_{post} \approx 10^{-3}$  is shown as dotted lines). The peak-to-peak metric differences at a  $BER_{post}$  of  $10^{-3}$  are 0.0064, 0.031, and 0.012 for  $BER_{pre}$ ,  $I_g$ , and  $I_a$ , resp.

Fig. 4 shows the relative post-FEC BER variation  $\Delta BER_{post} = \max\{BER_{post}\} / \min\{BER_{post}\}$  around  $BER_{post} \approx 10^{-3}$ . The variation in the fiber-optic channel is larger than in the Gaussian channel. The maximum  $\Delta BER_{post}$  are  $10^{5.2}$ ,  $10^{5.6}$ , and  $10^{2.6}$  for  $BER_{pre}$ ,  $I_g$ , and  $I_a$ , resp., over the fiber-optic channel, so  $I_a$  is clearly the best  $BER_{post}$  predictor ( $> 10^2$  times more accurate than the others). As for  $BER_{pre}$ , the accuracy is worse than  $I_a$  because of the steeper relationship



**Fig. 3:** ASI  $I_a$  or the post-FEC BER vs. launch power for some combinations of code rate and shaping.



**Fig. 4:** Relative post-FEC BER variation around  $BER_{post}$  of  $10^{-3}$  for various shaping and code rates.

between  $BER_{pre}$  and  $BER_{post}$ .

## Conclusions

We compared different metrics for post-FEC BER by numerical simulations. In the fiber-optic channel, though there are variations due to the lack of averaging or the burst errors, the characteristics are well correlated with that over the Gaussian channel. In the both channels, ASI is the best predictor. The ASI is more than 100 times more accurate than the pre-FEC BER or the normalized GMI at a post-FEC BER of around  $10^{-3}$  over the fiber-optic channel.

## References

- [1] F. Buchali et al., "Rate adaptation and ..." J. Lightw. Technol., vol. 34, no. 7, p. 1599 (2016).
- [2] F. Chang et al., "Forward error correction ...," IEEE Commun. Mag., vol. 48, no. 3, p. S48 (2010).
- [3] A. Alvarado et al., "Replacing the soft ...," J. Lightw. Technol., vol. 33, no. 20, p. 4338 (2015).
- [4] M. Ivanov et al., "On the information ...," IEEE Trans. Inf. Theory, vol. 62, no. 6, p. 3011 (2016).
- [5] T. Yoshida et al., "Performance metrics for systems with soft-decision FEC and probabilistic shaping," online: arXiv.org (2017).
- [6] ETSI, "Digital video broadcasting ..." EN 302 307-1, v1.4.1 (2014).
- [7] G. Böcherer et al., "Bandwidth efficient and ..." IEEE. Trans. Commun., vol. 63, no. 12, p. 4651 (2015).
- [8] P. Schulte et al., IEEE Trans. Inf. Theory, "Constant composition distribution ...," vol. 62, no. 1, p. 430 (2016).
- [9] T. Fehenberger et al., "Sensitivity gains by ..." IEEE Photon. Technol. Lett., vol. 28, no. 7, p. 786 (2016).
- [10] J. Cho et al., "Trans-Atlantic field ...," in Proc. OFC, paper Th5B.3 (2017).