Probabilistically-Shaped Coded Modulation with Hard Decision Decoding for Coherent Optical Systems

Probabilistically-Shaped Coded Modulation with Hard Decision Decoding for Coherent Optical Systems

Alireza Sheikh[(1)], Alexandre Graell i Amat[(1)], and Gianluigi Liva[(2)]
[(1)] Department of Electrical Engineering, Chalmers University of Technology, asheikh@chalmers.se
[(2)] Institute of Communications and Navigation of the German Aerospace Center (DLR)

Abstract *We consider probabilistic shaping to maximize the achievable information rate of coded modulation (CM) with hard decision decoding. The proposed scheme using binary staircase codes outperforms its uniform CM counterpart by more than 1.3 dB for 64-QAM and 5 bits/symbol.*

Introduction

Coded modulation (CM) schemes allow achieving high spectral efficiencies as required by the increasing demand for network capacity. By suitably choosing the probability of occurence of the constellation points, probabilistic shaping can significantly improve the efficiency of CM. It has been recently demonstrated that a probabilistically-shaped CM scheme using DVB-S2 codes with soft decision decoding (SDD) can operate within less than 1.1 dB of the additive white Gaussian noise (AWGN) capacity[1].

Hard decision decoding (HDD) can significantly reduce the complexity of the decoder compared to SDD. In particular, staircase codes[2], braided codes[3], and generalized product codes[4] with HDD have been proposed for high-speed fiber-optic communications, yielding very large net coding gains, yet with low decoding complexity. In this paper, we propose probabilistically-shaped CM with HDD. We consider a CM scheme using binary staircase codes with Bose-Chaudhuri-Hocquenghem (BCH) as component codes in combination with QAM modulation. We find the optimal distribution for the QAM constellation such that the bit-wise achievable information rate (AIR) of the CM system with HDD is maximized. We also design the parameters of the staircase code such that the requested information rate is achievable. By means of simulations, we show that the proposed CM scheme achieves significantly better performance compared to the baseline CM scheme with uniform distribution.

Preliminaries

We consider a discrete-time additive white Gaussian noise (AWGN) channel as an approximation of the fiber-optic channel (FOC). We also assume a 2^m-QAM constellation with average transmitted power $\mathbb{E}[|X_i|^2] \leq \mathsf{P}$, where $|\cdot|$ denotes the absolute value. Furthermore, a block-wise transmission system is considered, where $\boldsymbol{u}$ denotes the transmitted information block, and $\hat{\boldsymbol{u}}$ denotes the decoded information block. The block error probability of the system is defined as $\mathsf{P}_\mathsf{e} = \Pr(\boldsymbol{u} \neq \hat{\boldsymbol{u}})$. We consider a CM scheme with binary codes and HDD. Let X and $\hat{Y}$ be the random variables (RVs) corresponding to the channel input and hard detected output of the channel, respectively. Also, let B_X and $B_{\hat{Y}}$ be the RVs corresponding to the binary labels associated with X and $\hat{Y}$, respectively, using the binary reflected Gray code (BRGC) labeling. An AIR for the HDD system is given by the generalized mutual information (GMI), which can be computed as

$$\mathsf{I}_{\mathsf{HDD}}^{\mathsf{gmi}} = \sup_{s>0} \mathbb{E}_{B_X, B_{\hat{Y}}} \left[\log_2 \frac{q\big(B_X, B_{\hat{Y}}\big)^s}{\mathbb{E}_{B_{X'}}\left[q\big(B_{X'}, B_{\hat{Y}}\big)^s\right]} \right],$$

where s is the optimization parameter and $q\left(B_X, B_{\hat{Y}}\right)$ is the HDD metric computed based on the bit-wise Hamming metric, given as $q\left(B_X, B_{\hat{Y}}\right) = \varepsilon^{\mathsf{d_H}\left(B_X, B_{\hat{Y}}\right)}(1-\varepsilon)^{m-\mathsf{d_H}\left(B_X, B_{\hat{Y}}\right)}$, with $0 < \varepsilon < \frac{1}{2}$ and $\mathsf{d_H}\left(B_X, B_{\hat{Y}}\right)$ being the Hamming distance between binary labels B_X and $B_{\hat{Y}}$.

Probabilistic Shaping for the HDD System

We consider the Maxwell-Boltzmann distribution for the channel input X, which maximizes the bit rate of a discrete constellation for a given average energy[5]. Thus, the probability of selecting the constellation point $x \in \mathcal{X}$ is given by $p_X^\lambda(x) = A_\lambda e^{-\lambda|x|^2}$, where A_λ is a normalization factor and $\mathcal{X}$ is the QAM constellation alphabet. We assume that the QAM constellation is scaled with Λ to maintain the power constraint P. Therefore, the received symbol at the output of the channel is given by $Y = \Lambda X + Z$, with power constraint $\mathbb{E}[(\Lambda|X|)^2] = \mathsf{P}$, where Z is an independent and identically distributed (i.i.d.) complex Gaussian RV with zero mean and unit variance.

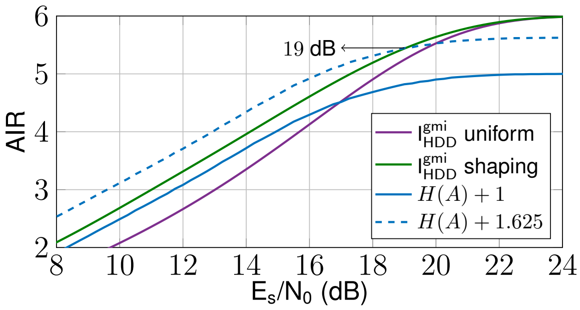

Fig. 1: AIRs (I^{gmi}_{HDD}) for 64-QAM and transmission rate of the designed CM scheme ($H(A)+2\gamma$) with $2\gamma=1$ and $2\gamma=1.625$.

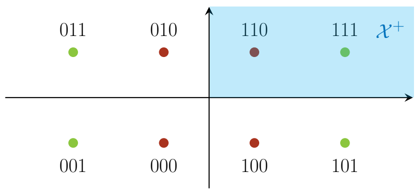

Fig. 2: 8-QAM constellation with BRGC labeling.

Tab. 1: Parameters of the designed staircase codes for $v=10$, $t=3$, and 64-QAM modulation

2γ	1.625	1.5313	1.4444	1.3023	1.1429	1
s	63	255	375	507	603	663
R_s	0.9375	0.9219	0.9074	0.8837	0.8571	0.8333

We find the shaping parameter λ and scaling Λ to maximize I^{gmi}_{HDD}, i.e., $(\Lambda^*,\lambda^*) = \underset{\Lambda,\lambda}{\text{argmax}}\ I^{gmi}_{HDD}$. To find (Λ^*,λ^*) for a given SNR, one can vary λ and find the corresponding Λ which satisfies the power constraint. Then, among all feasible pairs (Λ,λ), the one which maximizes I^{gmi}_{HDD} is selected. We denote by $p_X^{\lambda^*}$ the optimized distribution. In Fig. 1, we plot the AIR I^{gmi}_{HDD} for uniform and shaped distribution for 64-QAM constellation. As can be seen, the shaping significantly improves I^{gmi}_{HDD}. Let $\mathcal{X}^+$ be the set containing the 2^m-QAM constellation points with positive real and imaginary parts (see Fig. 2). We denote by A the RV which takes values on $\mathcal{X}^+$. Also, let S_R and S_I be the RVs correspodning to the sign of the real and imaginary part of the constellation points, respectively. It can be verified that $p_X^{\lambda^*}$ is symmetric with respect to the real and imaginary axis and that the RV X is represented uniquely by A, S_R, and S_I, i.e., $X=AS_RS_I$, with $p_A^{\lambda^*}(a)=4p_X^{\lambda^*}(a)$ and $p_{S_R}(1)=p_{S_I}(1)=p_{S_R}(-1)=p_{S_I}(-1)=\frac{1}{2}$. In Fig. 2, the constellation points with the same distribution are shown with the same color. Using the BRGC labeling, one can represent the points in the set $\mathcal{X}^+$ with $\log_2(2^m/4)=m-2$ bits, since the first 2 bits of the labeling of the points $x\in\mathcal{X}^+$ are the same. Using the mapping $0\mapsto -1$ and $1\mapsto +1$, the first and second bit of the labeling determine the sign of the real and the imaginary part of the constellation points, respectively.

Proposed Coded Modulation Scheme

The proposed CM scheme is shown in Fig. 3. Let $\boldsymbol{u}=(u_1,...,u_k)$ be the information block, of length k bits, generated by a uniform source. $\boldsymbol{u}$ is split into two vectors $\boldsymbol{u}^\mathsf{s}$ and $\boldsymbol{u}^\mathsf{a}$, of lengths $2\gamma n$ and $k-2\gamma n$, respectively. $\boldsymbol{u}^\mathsf{a}$ is used as the input of the shaping block (also called distribution matcher), shown with blue color in Fig. 3, where the output is from the set $\mathcal{X}^+$ with distribution $p_A^{\lambda^*}$. Here, we consider the constant composition distribution matcher[6]. The A_i's are transformed to length $m-2$-bit vectors $\boldsymbol{b}_i$ by the mapper Φ, which uses the BRGC labeling, and the resulting bits are concatenated using a parallel to serial converter. Therefore, $n(m-2)$ bits are contained in $\boldsymbol{b}_1,...,\boldsymbol{b}_n$.

We use staircase codes with binary systematic BCH codes as component codes. Let $\mathcal{C}$ be an $(n_\mathsf{c},k_\mathsf{c})$ shortened BCH code of (even) code length $n_\mathsf{c}=2^v-1-s$ and dimension $k_\mathsf{c}=2^v-vt-1-s$ constructed over the Galois field GF(2^v), where s is the shortening length and t the error correcting capability of the code. A shortened BCH code is thus completely specified by the triple (v,t,s). A staircase code with $(n_\mathsf{c},k_\mathsf{c})$ component codes is defined as the set of all $\frac{n_\mathsf{c}}{2}\times\frac{n_\mathsf{c}}{2}$ matrices $\boldsymbol{B}_\mathsf{i}$, $\mathsf{i}=1,2,\ldots$, such that each row of the matrix $[\boldsymbol{B}_{\mathsf{i}-1}^{\mathrm{T}},\boldsymbol{B}_\mathsf{i}]$ is a valid codeword of $\mathcal{C}$[2].

We set the rate of the staircase code to

$$R_\mathsf{s}=1-\frac{2(n_\mathsf{c}-k_\mathsf{c})}{n_\mathsf{c}}=\frac{m-2+2\gamma}{m}, \qquad (1)$$

where $1/2\le\gamma<1$ is a tuning parameter that can be used to select the rate of the staircase code and subsequently the spectral efficiency (SE) of the CM scheme. The sequences $\boldsymbol{b}_1,...,\boldsymbol{b}_n$ and $\boldsymbol{u}^\mathsf{s}$ form the information bits of one staircase block, with $n=\frac{(k_\mathsf{c}-n_\mathsf{c}/2)\cdot(n_\mathsf{c}/2)}{m-2+2\gamma}$. This specific value for n is due to the rate R_s in (1) and the structure of the staircase code where each component code is spread over two consecutive blocks. We select (ν,t,s) such that the number of parities in each staircase block is $n(2-2\gamma)$, which should be an integer number. Table 1 summarizes some of the designed code parameters for $\nu=10$ and $t=3$. One can easily show that the parity bits have a distribution that closely approximates the uniform one[1]. The uniform data bits $\boldsymbol{u}^\mathsf{s}$ are attached to the parity bits $p_1,...,p_{n(2-2\gamma)}$ and transformed to signs ($0\mapsto -1$ and $1\mapsto +1$) which are finally multiplied by the real and imaginary parts of $A_1,...,A_n$. The $X_1,...,X_n$ are then rescaled by Λ^* to maintain the power constraint before being sent over the FOC.

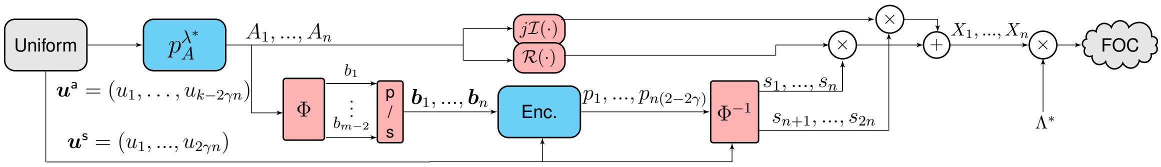

Fig. 3: Block diagram of the proposed CM scheme. $\mathcal{R}(\cdot)$ takes the real part while $\mathcal{I}(\cdot)$ takes the imaginary part.

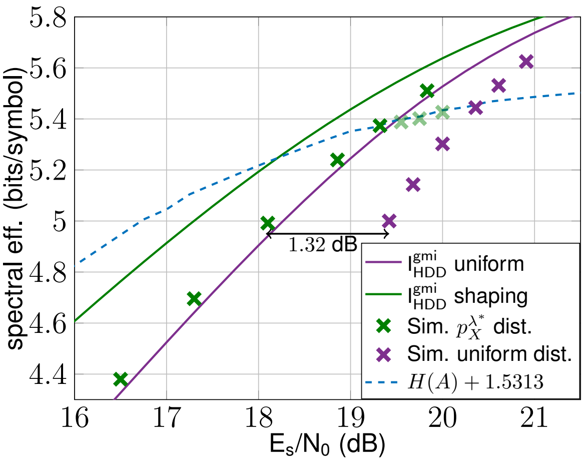

Fig. 4: AIRs (I_{HDD}^{gmi}) and simulation results for the proposed CM scheme with 64-QAM.

At the receiver, decoding of the staircase code is performed using bounded distance decoding and extrinsic message passing[3].

One can easily show that the transmission rate of the proposed CM scheme is $H(A) + 2\gamma$. Since vanishing error probability is only possible if $H(A) + 2\gamma < I_{HDD}^{gmi}$, by crossing the curves $H(A) + 2\gamma$ and I_{HDD}^{gmi}, one can find the feasible SNR region where the CM scheme can operate. In Fig. 1, for $2\gamma = 1$ all SNRs to the right of the curve $H(A) + 1$ are feasible while for $2\gamma = 1.625$ only the SNRs larger than 19 dB are feasible.

Simulation Results

We simulated the performance of the proposed CM scheme using the designed codes summarized in Table 1. We set the target block error probability to $P_e = 3 \times 10^{-3}$, corresponding to a bit error probability of roughly 10^{-7}, and by means of simulations we find the minimum required SNR such that the target P_e is met. This is shown with dark green crosses in Fig. 4. We remark that each point corresponds to a specific γ and it is on the curve $H(A) + 2\gamma$. This means that the points on the curve $H(A) + 2\gamma$ corresponding to larger SNRs can also meet the target performance. For example, for $2\gamma = 1.5313$, some other possible operational points are shown with faint green crosses. However, since by increasing the SNR $H(A) + 2\gamma$ becomes flat, at some point one needs to change γ (change the component code) to be as close as possible to the AIR I_{HDD}^{gmi}. We remark that one can vary the SE from 4.38 to 5.5 bits/symbol using a single code and decoder by simply changing the shaping distribution and shortening of the component code. Remarkably, our CM scheme with (finite-length) staircase codes achieves performances that are better than the AIRs (which assume ideal codes) of the CM scheme with uniform distribution. For the sake of comparison, we also simulated the performance of the system using the same codes with uniform distribution, i.e. (purple crosses). The proposed probabilistically-shaped CM scheme achieves significantly better performance compared to the uniform scheme. For a SE of 5 bits/symbol, the gain is 1.32 dB.

Conclusion

We proposed a CM scheme based on binary staircase codes with HDD for high-speed fiber-optic communications. The proposed scheme uses probabilistic shaping to maximize the AIR of the system. Remarkably, the performance of the proposed CM scheme with staircase codes is better than the AIR (i.e., assuming ideal codes) of the CM scheme with uniform distribution. Gains over 1 dB are observed compared to the performance of the baseline CM scheme with staircase codes and uniform distribution. Furthermore, the proposed CM scheme achieves performance within 0.5–1.2 dB of the AIR for a wide range of SEs using only a single staircase code, which greatly reduces the decoder complexity.

Acknowledgment

This work was financially supported by the Knut and Alice Wallenberg Foundation.

References

[1] G. Böcherer et al., "Bandwidth efficient and rate-matched low-density parity-check coded modulation," IEEE Trans. Commun., vol. 63, no. 12, pp. 4651–4665, Dec. 2015.
[2] B. P. Smith et al., "Staircase codes: FEC for 100 Gb/s OTN," J. Lightw. Technol., vol. 30, no. 1, pp. 110–117, Jan. 2012.
[3] Y. Y. Jian et al., "Iterative hard-decision decoding of braided BCH codes for high-speed optical communication," in Proc. GLOBECOM, 2013, pp. 2376–2381.
[4] C. Häger et al., "Density evolution for deterministic generalized product codes with higher-order modulation," in Proc. ISTC, Sep. 2016, pp. 236–240.
[5] F. R. Kschischang et al., "Optimal nonuniform signaling for gaussian channels," IEEE Tran. Inf. Theory, vol. 39, no. 3, pp. 913–929, May 1993.
[6] P. Schulte et al., "Constant composition distribution matching," IEEE Tran. Inf. Theory, vol. 62, no. 1, pp. 430–434, Jan. 2016.