A note on probabilistic and geometric shaping for the AWGN channel

Roberto Modonesi, Marco Dalai, Pierangelo Migliorati, and Riccardo Leonardi

Abstract—When considering coded modulation schemes for the AWGN channel, two main practical limitations prevent achieving channel capacity, namely the need to use a finite constellation and coding inefficiencies. Constellation shaping was given new impetus in recent works by Böcherer et al., which combined probabilistically shaped ASK constellations with LDPC coding. One open question is how far their results are from optimal solutions based on the use of finite constellations. We investigate this issue and show through experimental analysis that most of the loss is due to inefficient coding design and little benefits should be expected by adding two-dimensional geometric shaping.

Index Terms—AWGN channel, constellation shaping, power constraints.

I. INTRODUCTION

A well-known choice for the transmission of information in a communication system is the coded modulation paradigm [1], i.e., the combination of a coding block and a modulator that maps the codewords to symbols of a constellation, which are then sent to the channel. The performance can be raised in two ways. The first is the design and proper statistical use of the constellation, which brings an improvement called the *shaping gain*. The second is the adoption of an efficient code, causing an enhancement called the *coding gain*. The total gain is the combination of the two quantities, which can be considered separable and additive at high data rates [1]. This, in turn, means that the capacity of the communication channel in use may not be achieved if the constellation design is sub-optimal or the coding scheme is inefficient.

The constellation design can be improved by using constellation symbols with a non-uniform probability distribution (*probabilistic shaping* [2]) and/or non-uniformly spaced symbols (*geometric shaping* [3]). For the AWGN channel, the maximum shaping gain possible is 1.53 dB [4], [5], [6], [7, Section 4.3]. Both types of shaping were used in [8], which studies the convergence to the AWGN capacity as the constellation cardinality approaches infinity.

Recently, Böcherer et al. [2] proposed a novel constellation shaping for the AWGN channel. Their work raised the interest of the community because it provides a fine-tuning mechanism for the achievable rate and SNR parameters. This enables operating at $1 \cdot 10^{-3}$ frame error rate using length 64800 bit DVB-S2 LDPC codes, with a maximum gap to the channel capacity of 1.1 dB over a wide range of SNR values using Maxwell-Boltzmann distributions on equidistant ASK constellations.

When designing a coded modulation system it is important to know if the focus should be on the constellation design or on the choice of the coding algorithm. To evaluate possible improvements over the scheme presented in [2], it is important to appraise what is the impact of the constellation choice on the overall performance and whether the chosen constellation is used optimally. A reasonable question is what kind of constellation one should expect when the constellation design is done by both probabilistic and geometric shaping, and what margins of improvement one should expect. While [2] uses ASK modulation with Maxwell-Boltzmann probability distributions, we studied the margin of improvement, in terms of achievable input-output mutual information (MI), when adding geometric shaping in one and two dimensions, optimizing both positions and probabilities jointly. Our conclusion is that only little improvements are to be expected from constellation optimization and the focus should be on the coding scheme.

The paper is structured as follows. Section II shows the methods used to compute the optimal input distribution in the scalar and quadrature cases, while Section III presents the results and consequent discussions. The algorithms discussed here were implemented in MATLAB, using the built-in optimization function *fmincon*.

II. COMPUTATION OF THE OPTIMAL DISTRIBUTIONS

A. Scalar case with peak amplitude constraint

The discrete-time memory-less Scalar Gaussian Channel (SGC) is defined by the equation $y_i = x_i + n_i$, where y_i, x_i, and n_i denote the output, input, and noise of the communication channel at the (discretized) time instant i. The noise elements n_i are independent and identically distributed (i.i.d.) normal random variables that are independent of x_i.

Under the average power constraint $\mathbb{E}\left[|x|^2\right] \leq P_a$, the capacity is achieved by a Gaussian input, which induces a MI $\mathbf{I}(x; y) = \frac{1}{2}\log(1 + \text{SNR})$. Any modulation scheme uses a finite set of input values, i.e., a constellation, and this implies that such an input-output MI is not really achievable in such practical schemes. It is important then to re-set the reference of what is achievable and what is not under practically unavoidable system constraints. In this respect, we observe that the need for practical schemes to use a finite contellation is not really an inefficiency once one accepts the fact that practical schemes will never be able to use arbitrarily large input values. Indeed, under this assumption, Smith [9] proved that the distribution maximizing the input-output MI has finite support and one might then conceive a scheme using that support as the underlying constellation. So,

a reasonable reference for rates which could conceivably be achieved in practical systems can be established by considering the capacity of the AWGN channel under both average and peak power constraints. Computing the optimal support and probability mass function (pmf) of x, then requires solving a constrained optimization problem where the function to be maximized is $\mathbf{I}(x; y)$ and the power constraints are defined above. The rate of practical coding schemes should then be compared to the mutual-information achieved by such finite distribution.

In order to evaluate the efficiency of the different blocks of the method proposed in [2], we consider the same conditions of average power and peak input power and optimize the input support and distribution following the method proposed in [9].

B. Quadrature case with amplitude constraint

We also consider the discrete-time memory-less Quadrature Gaussian Channel (QGC) [10], defined by the equation

$$\underline{y}_i = \begin{bmatrix} y_{i_1} \\ y_{i_2} \end{bmatrix} = \underline{x}_i + \underline{n}_i = \begin{bmatrix} x_{i_1} \\ x_{i_2} \end{bmatrix} + \begin{bmatrix} n_{i_1} \\ n_{i_2} \end{bmatrix}. \quad (1)$$

Again i denotes the discretized time instant, whereas $\underline{x}$, $\underline{y}$, and $\underline{n}$ are complex random variables that denote the input, output, and noise of the QGC. The noise elements $\{\underline{n}_i\}$ are i.i.d. Gaussian random variables independent of $\{\underline{x}_i\}$ with components $\{n_{i_1}, n_{i_2}\}$ that are independent normal random variables. We consider for the input the peak and average power constraints P_p and P_a in the form $|\underline{x}_i|^2 \leq P_p$, $\forall i$ and $\mathbb{E}[|\underline{x}|^2] \leq P_a$. We consider (1) in its polar form [10]

$$\underline{x} = r\, e^{i\Theta} \quad \underline{y} = R\, e^{i\Psi} \quad r, R \geq 0 \quad \Theta, \Psi \in [-\pi, \pi) \quad (2)$$

since it is known that the distribution maximizing the input-output MI is supported on a finite number of radii, with uniform phase independent of the radius [10]. For every working point of interest in [2], the pmf of r is computed by doubling both power constraints to account for the 2-dimensionality of the problem (see Section II-A) and following the procedure described in [10]. The optimal distribution found has uniform phase and thus, contrarily to the scalar case, the MI achieved by this distribution is not a feasible target for practical modulation schemes which are constrained to use a finite constellation of points.

C. Phase quantization

One approach to determine a practical constellation is to quantize the support of the optimal uniform-phase input distribution. This process introduces the MI loss $\epsilon \triangleq |\mathbf{I}(x; y) - \mathbf{I}(x_Q; y_Q)|$, where the Q subscript denotes the phase quantization. The loss can be made as small as desired using a sufficiently fine quantization.

In order to test the quantizer proposed in [11] we set a maximum allowed ϵ, then we obtain a sufficiently high number of mass points ([11, eq. (13a)]) and their distribution over the circles of the optimal constellation ([11, eq. (16)]). The quantized constellation is computed by arranging the mass points on each circle in an equidistant manner. Finally, the actual MI loss is then computed

In our experiments, we considered here an acceptable MI loss of $\epsilon = \frac{1}{100}\mathbf{I}(x; y)$.

D. Actual number of required points

The procedure adopted in Section II-C is based on the results of [11] which provide a sufficient number of points that limits the MI loss to an acceptable value, but the actual number of points required may be much lower. In other words, the number of points derived in [11] actually allows for a much smaller value of ϵ than originally set. Also, when considering a finite number of usable points, quantizing the support of the optimal distributions may not be the optimal approach. So, we studied the optimal discrete input distributions that achieve the maximum of $\mathbf{I}(x; y)$ for a given cardinality of the support. The algorithm that computes the minimum number of points required and the corresponding distribution starts by considering 2 possible points and continues as follows:

- The candidate input distribution is obtained by maximizing $\mathbf{I}(x; y)$ subject to the peak and average power constraints.
- It checks if the information loss w.r.t the theoretically optimal one in Section II-B is higher than ϵ.
- If this happens, it increases the number of points by one and repeats the process, else it stops.

E. Circularly symmetric distributions

The optimal finite constellations found in the previous section have a geometry which renders them not very easy to deal with in terms of coding. The reference scheme [2] which we take as our reference work for our analysis is based on the idea of mapping information bits to channel input amplitudes by means of Constant Composition Distribution Matcher (CCDM) [12]. The signs of the inputs are partially taken from information bits and partially from the parity bits of an LDPC code. In the quadrature setting this corresponds to associating amplitudes given by a CCDM to a certain quadrant of the Cartesian plane and then index the four quadrants using either information or parity bits. In order to apply this scheme, a symmetry condition is then required on the support of the used constellation. So, we also computed the optimal finite support input distribution under the constraint that it be symmetric with respect to the coordinate axes. By calling the set of points of the target constellation as $\mathcal{X}$ we have then that $(x_1, x_2) \in \mathcal{X} \implies \{\pm x_1, \pm x_2\} \in \mathcal{X}$, $\forall x \in \mathcal{X}$ where $\{x_1, x_2\}$ denotes the coordinates of the points on the Cartesian plane. The algorithm used to compute the adjusted distribution starts by considering 1 mass point in each quadrant and works as follows:

- The power constraints are computed as per Section II-B.
- The candidate distribution is obtained by maximizing $\mathbf{I}(x; y)$ subject to the power constraints and the constraint of symmetry with respect to the coordinate axes.
- The MI loss is computed with respect to the MI obtained by using the optimal distribution in Section II-B.
- If the loss is higher than ϵ, it doubles the number of mass point in each quadrant and retries, otherwise it stops.
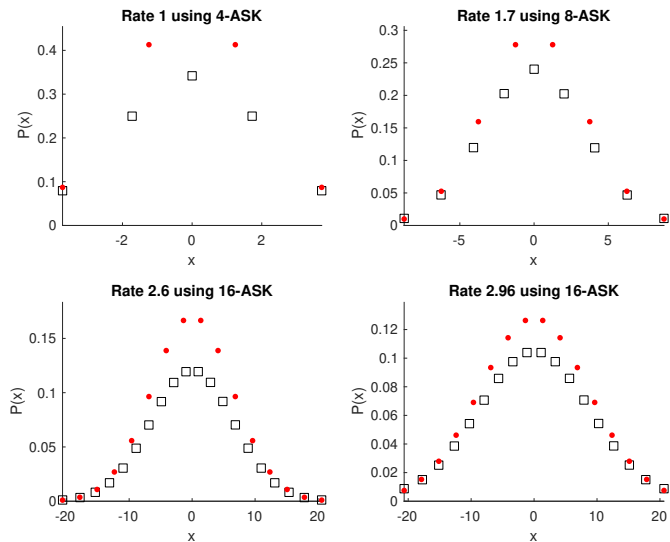
Fig. 1. Comparison between ASK (dots) constellations of [2] and corresponding constellations optimized as per Section II-A (squares). We also report the transmission rates obtained in [2].
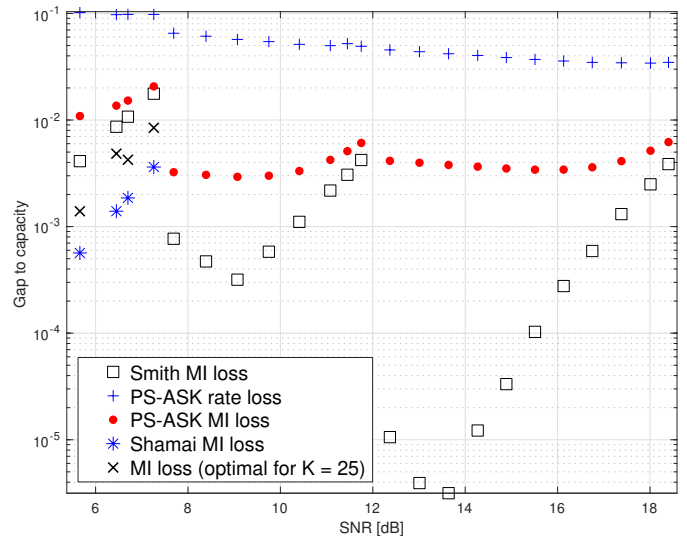

Fig. 2. Comparison of losses between Probabilistically Shaped ASK (PS-ASK) constellations of [2] and corresponding constellations optimized as per Sections II-A, II-B, and II-D.

III. RESULTS AND DISCUSSIONS

A. Results in the scalar case

We begin this section with a comparison between the shape of the constellations used in the reference scheme [2] and the shape of the constellations optimized as per Section II-A. This comparison shows what one should expect when the design is done by both probabilistic and geometric shaping. Figure 1 shows the comparison between the ASK distributions used in [2] (dots) and the optimal distributions computed as per Section II-A (squares). In particular, the figure shows one working point for the 4-ASK, one for the 8-ASK, and two for the 16-ASK cases for conciseness. We also report the rates achieved in [2] with block error probability near 10^{-3}. This section considers the working points of [2] that require up to 16-ASK constellations due to limited computational resources.

The optimized distributions have a slightly higher cardinality and are characterized by not equidistant mass points.

Figure 2 shows the losses in the two cases. The squares represent the loss of $\mathbf{I}(x; y)$ as per Section II-A. The plus signs represent the loss of transmission rates obtained in [2]. The dots are the loss of $\mathbf{I}(x; y)$ as per [2]. The remaining markers will be explained in section III-B. The data in the figure was divided by the Gaussian channel capacity for improved visibility. In a coded modulation system, the choice of the constellation lowers the channel capacity to the corresponding value of $\mathbf{I}(x; y)$ because of the sub-optimality of the chosen constellation. In [2] this MI loss has an average value of $1.1 \cdot 10^{-2}$ [bits]. The capacity is further lowered by losses due to the inefficiency of the coding block. When both losses are considered, the average rate loss as per [2] is $10^{-1} \left[\frac{\text{bits}}{\text{ch.use}}\right]$. Optimizing the input distribution as per Section II-A lowers the MI loss to the average value of $4.4 \cdot 10^{-3}$ [bits]. As shown in figure 2 most of the loss is due to the inefficiency of the coding block. The figure also shows that before the cardinality of the ASK constellations in [2] changes (from 4 to 8 at the

fourth working point and from 8 to 16 at the twelfth working point) the MI losses become more relevant. In [2] this behavior is due to the limited cardinality of the chosen constellations. When the input distribution is optimized as per Section II-A, the problem is the sub-optimal choice of the power constraints.

B. Results in the 2-dimensional case

As seen in figure 2, the losses assume more relevant values in the low SNR regime. It is thus interesting to observe if this behavior changes in the quadrature case. As such, figure 2 compares the losses (with respect to the channel capacity) in the scalar and quadrature cases in the low SNR regime. The asterisks represent the loss in the quadrature case as per Section II-B and the crosses represent the loss obtained when the constellation is optimized as per Section II-D. The data in the figure was divided by the respective channel capacity for improved visibility. As in the scalar case the cardinality of the optimized constellations is $K = 5$, the cardinality of the constellations optimized as per Section II-D is constrained to be K^2 to ensure a fair comparison. The figure shows that the constellations optimized as per Section II-B provide the lowest losses, while the constellations optimized as per Section II-D provide lower losses than the ones in the scalar cases, (note here that the peak power constraint is in 2D).

C. Results of the phase quantization

Using the phase quantization algorithm of Section II-C with a target maximal MI loss of $\epsilon = \frac{1}{100} \mathbf{I}(x; y)$ provides constellations that ensure acceptable losses. The cardinality of these constellations is greater than $1.6 \cdot 10^3$, which is shown to be two orders of magnitude higher than required in Section III-D. The proposed assignment of the quantized mass points [11, eq. (16)] must be slightly modified to assign points at circles with small radii.
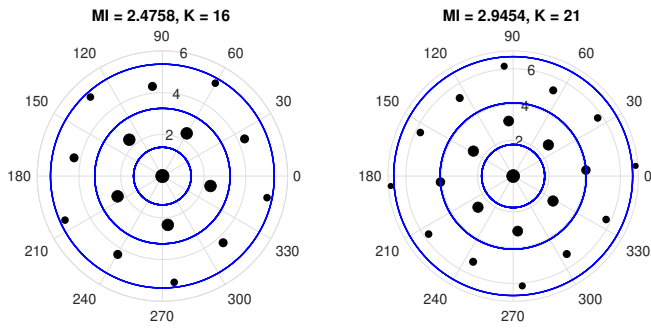
Fig. 3. Distributions with minimal number of points (bigger points are more probable). The circles represent the optimal constellations as per [10].

Fig. 4. Circularly symmetric distributions (bigger points are more probable). The circles represent the optimal constellations as per [10].

D. Results on the actual number of required points

The algorithms presented in Section II-D and Section II-E were used over the working points of [2] up to an ASK cardinality of 8 for limited computational capabilities. The actual number of required points ranges from 14 to 17 when the corresponding QAM cardinality is 16. When the cardinality is 64 the values range from 19 to 43 as the power constraints become more relaxed. Figure 3 shows the optimal distributions in two cases: the first one corresponds to an achievable rate [2] (by extension to the 2-dimensional case) of 2.26 $\left[\frac{\text{bits}}{\text{ch.use}}\right]$ while the second one corresponds to a rate of 2.8 $\left[\frac{\text{bits}}{\text{ch.use}}\right]$.

When the constellation cardinality of [2] is 16 ($2 \leq$ rate $\leq$ 2.4) the number of points required increases slowly with the rate and the mass points usually don't intersect the theoretical circles. When the constellation cardinality of [2] is 64 ($2.6 \leq$ rate $\leq$ 3.8), the distributions become more similar to phase quantizations of the optimal ones as the rate increases.

E. Results on the circularly symmetric distributions

At lower rates (rate $\leq$ 2.4) the number of points required to get an acceptable MI loss ranges from 16 to 32. The first constellation resembles a cross-QAM with the lower energy points being more probable than the others, while the remaining three constellations are similar to phase quantizations of the optimal ones. At middle rates ($2.6 \leq$ rate $\leq$ 3.2) the number of points is halved w.r.t. the QAM cardinalities ($K = 32$) and the constellations are similar to cross-QAM, while at higher rates ($3.4 \leq$ rate $\leq$ 3.8) the number of points is the same as the QAM cardinalities ($K = 64$) and the constellations become more similar to phase quantizations of the optimal ones. Figure 4 shows two cases: in the first, the distribution (the discrete set of points) resembles a probabilistically shaped 16-QAM while in the second case it resembles a 32 cross-QAM.

IV. CONCLUSIONS

We considered the AWGN channel subject to peak and average power constraints. We studied the performance of the method proposed in [2] investigating the loss due to coding and the loss due to the constellation choice and probability shaping. We then studied practical constellations obtained with different methods, focusing on the required cardinality. Our results provide evidence that most of the loss in [2] is to be attributed to coding and little margin of performance are
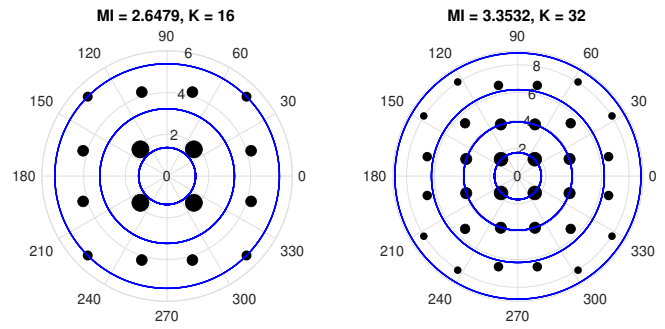
to be expected from further optimization of the probabilistic and geometric shaping of constellations. Future work should be divided in three parts. A possible follow-up could be the extension of the results provided by [9] and [10] to optimize the constellation under the assumption of a bit-metric decoder. Then for the QGC it would be interesting to find a closed-form expression for the number of mass points of the constellation required to achieve capacity as a function of the power constraints. Finally, specific bit mapping and coding should be tested on circularly symmetric distributions as to evaluate the potentially achievable results with a scheme similar to what is proposed in [2].

REFERENCES

[1] G. D. Forney. Trellis shaping. *IEEE Transactions on Information Theory*, 38(2):281–300, March 1992.

[2] G. Böcherer, Fabian Steiner, and Patrick Schulte. Bandwidth efficient and rate-matched low-density parity-check coded modulation. *IEEE Transactions on Communications*, 63(12):4651–4665, 2015.

[3] F. Steiner and G. Böcherer. Comparison of geometric and probabilistic shaping with application to atsc 3.0. In *SCC 2017; 11th International ITG Conference on Systems, Communications and Coding*, pages 1–6, Feb 2017.

[4] A. R. Calderbank and L. H. Ozarow. Nonequiprobable signaling on the gaussian channel. *IEEE Transactions on Information Theory*, 36(4):726–740, July 1990.

[5] G. D. Forney, R Gallager, G Lang, F Longstaff, and S Qureshi. Efficient modulation for band-limited channels. *IEEE Journal on Selected Areas in Communications*, 2(5):632–647, 1984.

[6] F. R. Kschischang and S. Pasupathy. Optimal nonuniform signaling for gaussian channels. *IEEE Transactions on Information Theory*, 39(3):913–929, May 1993.

[7] G. Böcherer. Principles of coded modulation. *Habilitation Thesis, Technische Universität München*, 2018. [Online]. Available: http://www.georg-boecherer.de/bocherer2018principles.pdf.

[8] Yihong Wu and Sergio Verdú. The impact of constellation cardinality on gaussian channel capacity. In *2010 48th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, pages 620–628. IEEE, 2010.

[9] Joel G Smith. The information capacity of amplitude-and variance-constrained scalar gaussian channels. *Information and Control*, 18(3):203–219, 1971.

[10] Shlomo Shamai and Israel Bar-David. The capacity of average and peak-power-limited quadrature gaussian channels. *IEEE Transactions on Information Theory*, 41(4):1060–1071, 1995.

[11] Wasim Huleihel, Ziv Goldfeld, Tobias Koch, Mokshay Madiman, and Muriel Médard. Design of discrete constellations for peak-power-limited complex gaussian channels. In *2018 IEEE International Symposium on Information Theory (ISIT)*, pages 556–560. IEEE, 2018.

[12] P. Schulte and G. Böcherer. Constant composition distribution matching. *IEEE Transactions on Information Theory*, 62(1):430–434, Jan 2016.