On Constellation Shaping for Short Block Lengths

Y.C. Gültekin W.J. van Houtum F.M.J. Willems
Information and Communication Theory Lab
Signal Processing Systems Group
Department of Electrical Engineering, TU/e, The Netherlands
{y.c.g.gultekin, w.j.v.houtum, f.m.j.willems}@tue.nl

Abstract

Gaussian channel inputs are required to achieve the capacity of additive white Gaussian noise (AWGN) channels. Equivalently, the n-dimensional constellation boundary must be an n-sphere. In this work, constellation shaping is discussed for short block lengths. Two different approaches are considered: Sphere shaping and constant composition distribution matching (CCDM). It is shown that both achieve the maximum rate and generate Maxwell-Boltzmann (MB) distributed inputs. However sphere shaping achieves this maximum faster than CCDM and performs more efficiently in the short block length regime. This is shown by computing the finite-length rate losses. Then the analysis is justified by numerical simulations employing low-density parity-check (LDPC) codes of the IEEE 802.11 standard.

1 Introduction

We consider the transmission of information X over a one-dimensional AWGN channel under the average power constraint P. The capacity-achieving input distribution $P(x)$ for this channel is the zero-mean Gaussian with variance P. The loss in maximum achievable information rate (AIR) resulting from using a uniform $P(x)$ is called the shaping gap and is 0.255 bits per real dimension asymptotically in signal-to-noise ratio (SNR) and block length n. This loss can also be interpreted as the increase in required SNR to achieve a certain rate R and is 1.53 dB asymptotically in R and n. Fig. 1 illustrates this loss by plotting channel capacity and the maximum AIR for a uniform input distribution over $[-\sqrt{3P}, \sqrt{3P}]$.

On top of the shaping gap, there is also a performance degradation caused from the discrete nature of the practically used channel inputs X. Fig. 1 also illustrates this by plotting the mutual information (MI) between the channel input X and channel output Y for 2^m-ary amplitude shift keying (ASK) alphabets

$$\mathcal{X} = \{\pm 1, \pm 3, \cdots, \pm (2^m - 1)\}. \tag{1}$$

These curves stay close to the uniform capacity until $R = (m-1)$ and then converge to m asymptotically in SNR. In this paper, we consider ASK based transmission schemes while investigating several constellation shaping methodologies.

Constellation shaping can be defined as the optimization of channel input X with the purpose of decreasing the required average transmit power for a given target error probability. This topic is investigated from many perspectives ever since Shannon determined the capacity-achieving distribution for AWGN channels in his celebrated paper. For a detailed survey on signal shaping, see [1].

In this paper, two fundamentally different approaches for shaping will be examined. These are the recently-introduced constant composition distribution matching (CCDM) and the well-investigated sphere shaping. In Sec. 2, these methods are explained. In

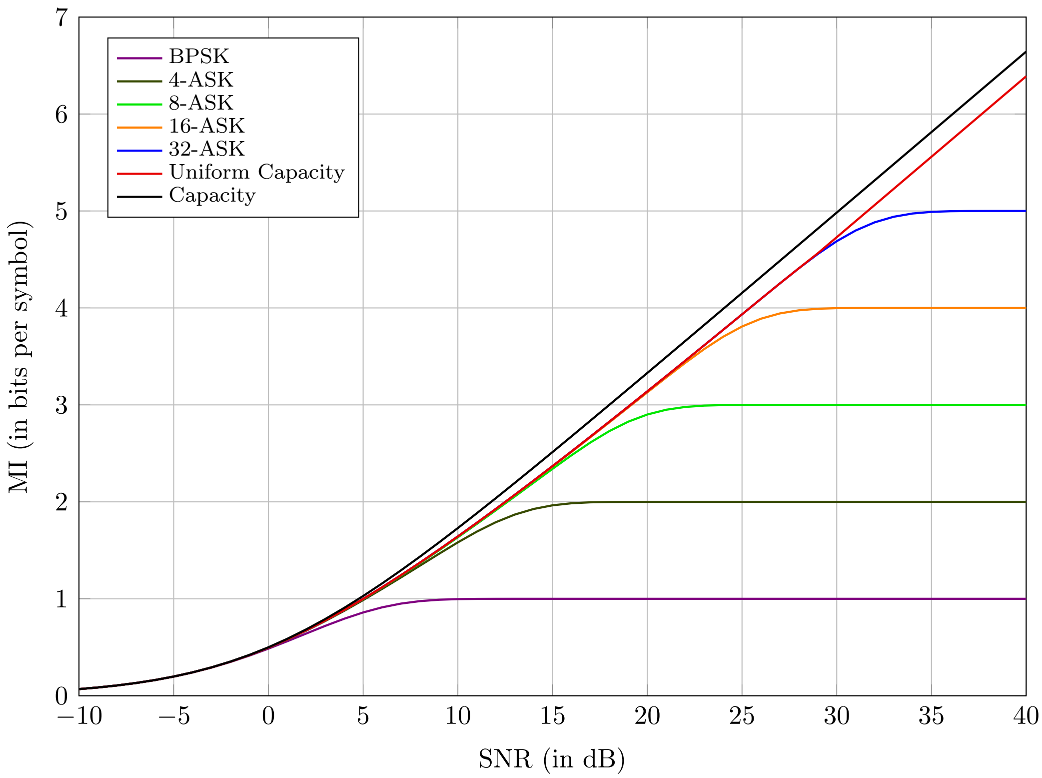

Figure 1: MI between channel input and output employing 2^m-ASK constellations for $m = 1, 2, \cdots, 5$ over AWGN channel. The Shannon capacity and the uniform capacity, i.e., the maximum AIR using uniform inputs, are also presented. The difference between the last two is 1.53 dB asymptotically, i.e., 0.255 bits per real dimension.

Sec. 3, an information-theoretical study is presented showing the asymptotic optimality of them. In Sec. 4, numerical results are presented to further illustrate the difference between the shaping approaches before the conclusion.

2 Shaping Approaches

Taxonomically, constellation shaping approaches are classified into two groups. The first one is probabilistic shaping (PS) where elements of $\mathcal{X}$ are employed with non-uniform probabilities. The second one is geometric shaping (GS) where positions of low-dimensional constellation points are optimized [1, Sec. 4.5]. We believe this categorization does not uncover the importance of sphere shaping where the boundary of the multidimensional constellation is structured in a way that will improve the performance. Although this again leads to non-uniform utilization of channel symbols as in PS, the indirect nature of accomplishing this motivates us to consider sphere shaping separately.

An information theoretically elegant way of constellation shaping is to first determine the capacity-achieving channel input distribution and then to obtain inputs with this distribution. Using channel input symbols with non-uniform probabilities according to a pre-defined distribution is called PS in this work. Recently, the concept of distribution matching (DM) is proposed to realize PS [2]. DM refers to any procedure that creates a non-uniform distribution that approximates, i.e., matches, a desired distribution. There are multiple ways of implementing distribution matchers in the literature: Variable-to-fixed length prefix-free coding [2], arithmetic coding [3] etc. The approach in [3] is called CCDM and is combined with channel coding in the probabilistic amplitude shaping (PAS) framework in [4]. CCDM attracted a lot of attention especially in the optical communication society due to its high efficiency at large block lengths and the ability to create a fine granularity for the transmission rate, or equivalently, reach. We will use CCDM to represent PS here and discuss its advantages, disadvantages and efficiency for short to medium block lengths n.

From another perspective, the shaping gap can also be closed by sphere-shaping the multidimensional constellation boundary. As shown in [1] and [5], if an n-sphere is used as the boundary of the n-dimensional 2^m-ASK constellation[*] $\mathcal{X}^n$, the induced distribution on any low-dimensional constituent constellation approaches to a Gaussian asymptotically in n and m. Thus, the objective of transmitting Gaussian inputs can also be achieved by n-sphere shaping. Furthermore, achieving this goal by imposing a sphere constraint on an n-dimensional lattice might be more elementary from the channel coding perspective.

Next CCDM and n-sphere shaping will be explained.

2.1 Constant Composition Distribution Matching

Following the approach and notation in [3], we factorize $\mathcal{X}$ as $\mathcal{X} = \mathcal{A} \times \mathcal{S}$ where

$$\mathcal{A} = \{1, 3, \cdots, (2^m - 1)\}, \tag{2}$$

$$\mathcal{S} = \{-1, 1\}, \tag{3}$$

and limit our focus to n-amplitude sequences $A^n \in \mathcal{A}^n$.

Let P_A indicate a discrete distribution over the amplitude alphabet $\mathcal{A}$. The n-type distribution $P_{\bar{A}}$ which minimizes[†] $\mathbb{D}(P_A || P_{\bar{A}})$ is used to determine the composition $\#(a)$ as

$$\#(a) = nP_{\bar{A}}(a) \text{ for } a \in \mathcal{A}. \tag{4}$$

Then in a similar way to [7], arithmetic coding is utilized in [3] to implement a matcher that indexes all possible n-tuples $a_1 a_2 \cdots a_n$ having the same composition, i.e., constant composition. The functional diagram of such a matcher is given in Fig. 2. The rate of this matcher which indexes n-sequences with k bits is $R_{cc} = k/n$ bits per symbol.

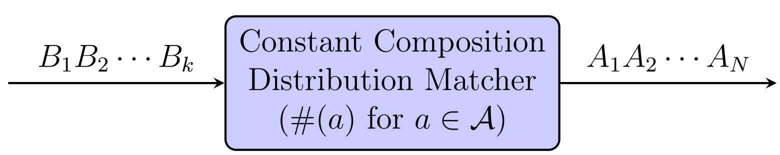

Figure 2: Constant composition distribution matcher. It maps k uniform bits (i.e., an index) to n amplitudes having a fixed composition $\#(a)$ for $a \in \mathcal{A}$. The mapping is invertible. Arithmetic coding is employed.

The distribution of these sequences at the output of the matcher is indicated by $P_{\tilde{A}^n}$. Geometrically, these sequences are located on the surface[‡] of an n-sphere of radius

---

[*]Here the term 'n-dimensional ASK constellation' is used to indicate the n-fold Cartesian product of $\mathcal{X}$.

[†]For the optimum way of computing $P_{\bar{A}}$, see [6].

[‡]Note that there are multiple compositions that lead sequences to the same surface.

$r = \sqrt{E_{cc}}$ where

$$E_{cc} = \sum_{a \in \mathcal{A}} \#(a) a^2, \tag{5}$$

is the sequence energy. It is shown in [3] that the informational divergence between the desired and the output distributions

$$\mathbb{D}(P_{\tilde{A}^n} || P_A^n) = \mathbb{D}(P_{\tilde{A}^n} || P_{\bar{A}}^n) + n\mathbb{D}(P_{\bar{A}} || P_A), \tag{6}$$

approaches zero for $n \to \infty$ which is equivalent to say for the rate of the constant composition code that

$$\lim_{n \to \infty} R_{cc} = \mathbb{H}\left(P_A\right). \tag{7}$$

Here it is assumed that the n-sequences are employed with equal probability. Note that letting $\mathcal{S}_{cc}$ indicate the set of these constant composition sequences, the rate can be written as

$$R_{cc} \triangleq \frac{\lfloor \log_2\left(|\mathcal{S}_{cc}|\right) \rfloor}{n} \text{ bits per symbol.} \tag{8}$$

Finally in [4], MB distributions are used as P_A to close the shaping gap relying on the fact that they have the maximum entropy for a fixed second moment, i.e., the variance[§] in this case [8].

The main advantages of CCDM are the asymptotic optimality and the virtual possibility of transmitting any rate by playing with P_A which can be invaluable in optical communications concerning the reach. On the other hand the disadvantages are the very long block length requirement to perform efficiently and inability to be parallelized. Note that the last is due to the use of arithmetic coding.

2.2 n-Sphere Shaping

Motivated by the asymptotic duality between n-sphere and Gaussian distribution, n-sphere shaping is the procedure of putting a maximum-energy constraint (i.e., the sphere constraint) on the possible n-sequences. The set $\mathcal{S}_\circ$ of these sequences can be defined as

$$\mathcal{S}_\circ = \left\{ a_1 a_2 \cdots a_n \middle| \sum_{i=1}^{n} a_i^2 \leq E_{\max} \right\}, \tag{9}$$

where $E_{\max}$ is the maximum sequence energy.

In the literature, the fact that a sphere constraint leads to a Gaussian distribution is proved using continuous approximation. Although there is no closed-form expression for the distribution that maximizes the AIR constrained by an ASK constellation, the reasoning in the continuous domain is extended to discrete domain somewhat pragmatically. Though in a coding setup, this makes sense since a shaping code can easily be combined with a (systematic) error-correcting code. In Sec. 3, we will show that the sphere constraint induces a MB distribution asymptotically when imposed on an n-dimensional ASK lattice.

[§]Here we implicitly assume that the distribution of signs will be uniform and $\mathcal{X}$ will have zero mean.

2.2.1 Enumerative Sphere Shaping

Although it is not necessary to specify an encoding strategy for n-sphere shaping to analyze its performance, here we outline enumerative sphere shaping (ESS) for the sake of completeness. Proposed in [9], ESS specifies efficient encoding and decoding algorithms to index energy-bounded sequences. These sequences are sorted lexicographically. In the context of this work, an enumerative shaper can be regarded as a black box, see Fig. 3. For a detailed discussion, see [5] and [9].

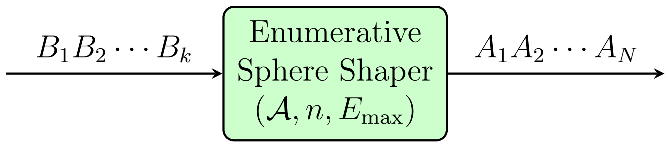

Figure 3: Enumerative sphere shaper. It is an invertible mapping from an index (i.e., k bits) to n amplitudes. All possible sequences have an energy no greater than $E_{\max}$.

Geometrically, energy bounded sequences are located in and on the surface of an n-sphere of radius $r = \sqrt{E_{\max}}$. The rate of this code is $R_\circ = k/n$ bits per symbol given that $E_{\max}$ is large enough to enclose more than $2^{nR_\circ}$ n-sequences. This rate can also be written as

$$R_\circ \triangleq \frac{\lfloor \log_2 (|\mathcal{S}_\circ|) \rfloor}{n} \text{ bits per symbol.} \tag{10}$$

Recently in [5], ESS is combined with non-systematic convolutional coding to improve the performance of IEEE 802.11 for $n = 96$.

Here we note that two different addressing algorithms for sphere shaping are given in [10]. The first algorithm is very similar to ESS where the only difference is sequences being sorted with respect to the n-dimensional shell that they are located on. The second algorithm is the well-known shell mapping which is introduced in [11] and included in the V.34 modem standard for $n = 16$ [12].

In the next section, we will investigate the asymptotic properties of constant composition and sphere codes.

3 Information-Theoretical Analysis

Let $\mathcal{C}$ be a code which consists of amplitude sequences $\boldsymbol{a}_k = a_{k,1} a_{k,2} \cdots a_{k,n}$ for $k = 1, 2, \cdots, L$. Here L indicates the number of codewords of length n in the code. All codewords occur with probability $1/L$.

The operational rate of such a code is $\log_2(L)/n$ bits per symbol and its operational average symbol energy is $\frac{1}{L}\sum_{k=1}^{L} \frac{1}{n} \sum_{i=1}^{n} a_{k,i}^2$.

Definition 3.1. (Achievability) The rate-energy pair (R, E) is called achievable if for each $\epsilon > 0$, for all n large enough, there exists a code with operational rate and operational average symbol energy satisfying

$$\frac{\log_2(L)}{n} \geq R - \epsilon, \tag{11}$$

$$\frac{1}{L}\sum_{k=1}^{L} \frac{1}{n} \sum_{i=1}^{n} a_{k,i}^2 \leq E + \epsilon. \tag{12}$$

Finally we define the rate-energy function as follows:

$$R(E) \stackrel{\Delta}{=} \max\{R : (R, E) \text{ is achievable}\}. \tag{13}$$

Theorem 1. The maximum achievable rate for average symbol energy E is

$$R(E) = \max_{A:\mathbb{E}[A^2]\leq E} \mathbb{H}(A).$$

The proof consists of a converse part and the corresponding achievability proof.

3.1 Converse

Consider a code. Assume that amplitude A_i for $i = 1, 2, \cdots, n$, has marginal distribution P_i generated[¶] by the uniform distribution over the codewords. Now the operational rate can be upper-bounded as

$$\begin{aligned} \frac{\log_2(L)}{n} &= \frac{1}{n}\mathbb{H}\left(A_1 A_2 \cdots A_n\right) = \frac{1}{n}\sum_{i=1}^{n} \mathbb{H}(A_i|A_{i-1}, \cdots, A_1) \\ &\stackrel{(a)}{\leq} \frac{1}{n}\sum_{i=1}^{n} \mathbb{H}\left(A_i\right) \stackrel{(b)}{\leq} \mathbb{H}(\tilde{A}), \end{aligned} \tag{14}$$

where (a) follows from the fact that conditioning cannot increase entropy. If we say that $\tilde{A}$ is a random variable over alphabet $\mathcal{A}$ with distribution

$$\tilde{P}(a) = \frac{1}{n}\sum_{i=1}^{n} P_i(a), \tag{15}$$

then (b) is due to the convexity of entropy.

Next we observe that

$$\frac{1}{L}\sum_{k=1}^{L}\frac{1}{n}\sum_{i=1}^{n} a_{k,i}^2 = \frac{1}{n}\sum_{i=1}^{n}\sum_{a\in\mathcal{A}} P_i(a)a^2 = \sum_{a\in\mathcal{A}} \tilde{P}(a)a^2 = \mathbb{E}[\tilde{A}^2].$$

We now conclude that for an achievable rate-energy pair (R, E), for all $\epsilon > 0$ and all large enough n, there exists a random variable A over $\mathcal{A}$ such that both

$$R \leq \frac{\log_2(L)}{n} + \epsilon \leq \mathbb{H}(A) + \epsilon, \tag{16}$$

$$E \geq \frac{1}{L}\sum_{k=1}^{L}\frac{1}{n}\sum_{i=1}^{n} a_{k,i}^2 - \epsilon = \mathbb{E}[A^2] - \epsilon. \tag{17}$$

If we let $\epsilon \downarrow 0$ we obtain that

$$R(E) \leq \max_{A:\mathbb{E}[A^2]\leq E} \mathbb{H}(A). \tag{18}$$

[¶]More precisely, for $a \in \mathcal{A}$, $P_i(a)$ is te number of codewords $\boldsymbol{a}_k$ for which $a_{k,i} = a$ divided by L.

3.2 Achievability Part

3.2.1 Achievability Based on Constant Composition Codes

Fix an energy E and assume that random variable A^* maximizes the entropy $\mathbb{H}(A^*)$ while satisfying the energy constraint $\mathbb{E}[(A^*)^2] \leq E$. Denote by $\{P^*(a), a \in \mathcal{A}\}$ the (MB) distribution corresponding to this random variable. For all large enough n, we now take a composition $\#(a)$ that satisfies

$$|\#(a) - nP^*(a)| \leq 1, \tag{19}$$

where $\#(a) \geq 0$ for all $a \in \mathcal{A}$, and $\sum_{a \in \mathcal{A}} \#(a) = n$.

It can be shown that the probabilities $P^*(a) > 0$, see [8, Example 11.2.3]. Therefore the normalized composition $\{\#(a)/n, a \in \mathcal{A}\}$ approaches entropy $\mathbb{H}(P^*)$ for increasing n.

Now for fixed n, consider a code consisting of all sequences having the composition $\{\#(a), a \in \mathcal{A}\}$. It can be shown that the operational rate of this constant composition code approaches the entropy $H(\#(A)/n)$ of the normalized composition for increasing n, see again [8, Example 11.2.3], where Stirling approximation is used.

We conclude that the operational rate of the constant composition code approaches the entropy of the normalized composition, which approaches entropy $\mathbb{H}(P^*)$, for n large. Therefore $R(E) = \max_{A:\mathbb{E}[A^2]\leq E} \mathbb{H}(A)$ is achievable for all E.

3.3 Optimality of Sphere Codes

Definition 3.2. (Sphere Code) A code is a sphere code if there exist no sequences, not in the code, with energy smaller than the energy of a codeword.

Theorem 2. For each code with operational rate R and energy E, there is a sphere code with operational rate R_o and operational symbol energy E_o such that

$$R_o = R \text{ and } E_o \leq E. \tag{20}$$

Proof. Just replace codewords by sequences outside the code with lower energy until the code is a sphere code. □

Theorem 2 along with the optimality of constant composition codes leads to the conclusion that sphere codes achieve the maximum rate as well.

Note that the code $\mathcal{S}_\circ$ defined before in Section 2.2 is a sphere code. The enumerative sphere code with rate as in (10) need not be a sphere code since sequences are sorted lexicographically and the ones with index 2^{nR} or more are not used. Observe that in the first algorithm in [10], sequences with largest index also have the largest energy and therefore the remaining sequences form a sphere code.

3.4 Maxwell-Boltzmann Distribution and Comparison

The distribution that achieves maximum entropy under an energy constraint is called Maxwell-Boltzmann distribution, see [8]. It is easy to show that the maximum entropy distribution is unique. This follows directly from the strict convexity of the entropy function. Since sphere codes result in maximum entropy under an energy constraint, the corresponding average marginal distribution (15) approaches the MB distribution.

To show the superiority of sphere codes over constant composition codes for short block lengths, we compute the finite-length rate loss

$$R_{\text{loss}} = H(A) - R, \tag{21}$$

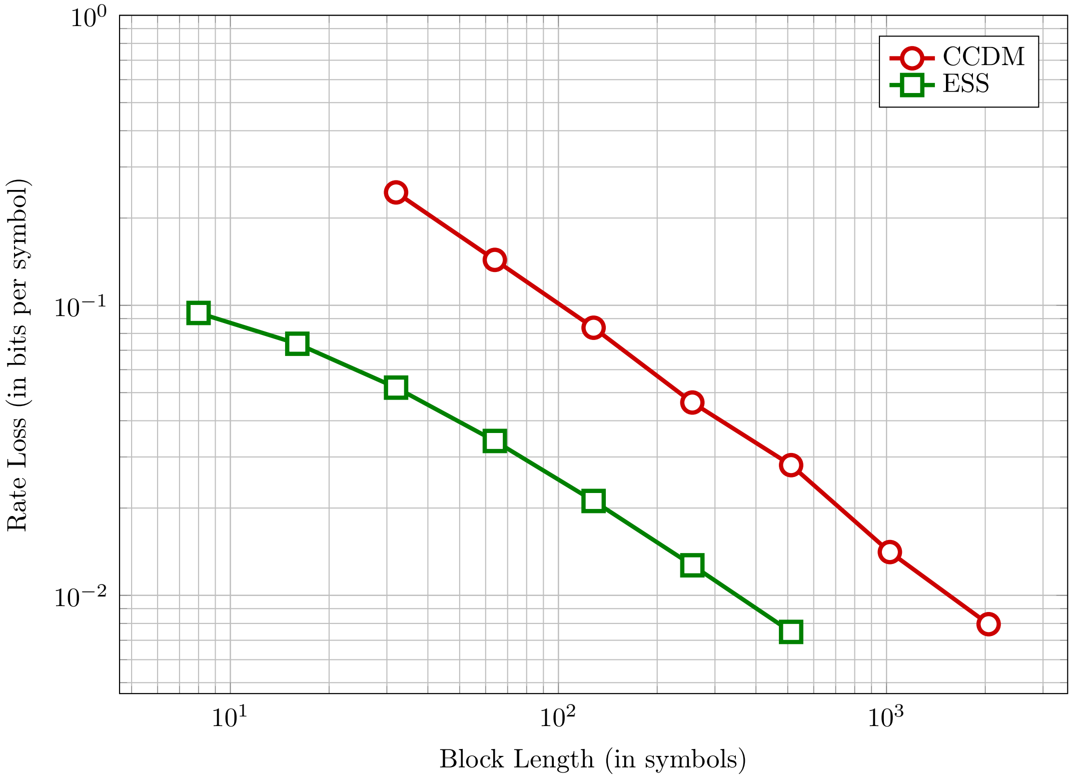

Figure 4: Block length (in symbols) vs. finite-length rate losses (in bits per symbol) of constant composition and sphere codes. The target rate is $R = 1.75$ for $\mathcal{A} = \{1, 3, 5, 7\}$.

where A is MB distributed and $\mathbb{E}[A^2]$ is equal to the average energy of the code used to achieve R. The results are presented in Fig. 4.

Here we fix the target rate $R = 1.75$, and find the composition $\#(a)$ and $E_{\max}$ that achieve R for constant composition and sphere codes respectively. It appears that for a target rate loss of 0.1 and 0.01 bits per symbol, constant composition codes require approximately 10 and 5 times larger block lengths than sphere codes, respectively. Furthermore, due to their definition, see Definition 3.2, sphere codes have the smallest block length requirement for a target rate loss.

3.5 Notes

We note that one counter argument may be the following: The shaping gain, i.e., the gain in average energy or in rate, does not necessarily imply a similar gain in AIR or an SNR improvement for a given error probability. Furthermore, utilization of constant composition codewords may enable better decoding performance than a sphere code in some cases. As an example, the constant composition property is exploited by a type check in successive cancellation list decoding of polar codes in [13].

A second note concerns the work presented in [14]. Here the shell mapping which is also an indexing method for n-sphere shaping is compared with CCDM. There are two fundamental differences between shell mapping and ESS: First, shell mapping sorts sequences with respect to their energies, i.e., the n-dimensional shell that they are located on, instead of lexicographical ordering. Second, the algorithm employs the divide and conquer principle (which requires multiplications) instead of operating in a sequential manner.

Finally, a bridge between sphere shaping and CCDM, namely partition-based distribution matching (PBDM), is established in [15]. In this work, the constant composition constraint is relaxed by employing multiple compositions having the desired composition as the ensemble average. To put it simply, instead of considering a single shell, multiple nested shells are utilized. To select the corresponding composition, a variable-length prefix of the binary input is used which can be considered as a disadvantage. Although not as efficient as a sphere code, PBDM also achieves the maximum rate faster than CCDM.

4 Numerical Results

Monte Carlo simulations are used to compare the performance of constant composition and sphere codes in the short to medium block length regime. For CCDM simulations, PAS scheme is realized as in [4]. For sphere shaping simulations, the matcher in the PAS scheme is replaced by an enumerative sphere shaper. As the channel code, systematic LDPC codes of IEEE 802.11 [16] are employed with two different codeword lengths, i.e., $n_c = \{648, 1944\}$ bits. Note that the 2^m-ASK demapper on the receiver side assumes that the symbols are independent and identically distributed with either the n-type distribution of the constant composition code or the average marginal distribution of the sphere code. To achieve the target rate $R = 2.67$ based on 16-ASK, uniform transmission is simulated with the rate $r_c = 2/3$ code whereas shaped transmissions require $r_c = 3/4$.

Table 1: Parameters for Shaping

Approach	n	#(a) or $E_{\max}$	# seq.	E	Gain (in dB)
CCDM	162	(34, 32, 28, 23, 18, 13, 9, 5)	$2^{432.06}$	48.31	0.44
	486	(112, 103, 88, 69, 50, 33, 20, 11)	$2^{1296.15}$	42.22	1.02
ESS	162	796	$2^{432.13}$	39.74	1.29
	486	2326	$2^{1296.02}$	39.10	1.35

To obtain 2^{nR} n-sequences, the compositions and $E_{\max}$ values given in Table 1 are used for CCDM and ESS, respectively. Energy per symbol and shaping gain values of these schemes are also given in the same table. The shaping gain is computed with respect to the average energy equation $(2^{2H(X)} - 1)/3$ of uniform ASK constellations which is 53.42 for this target rate[‖].

In Fig. 5, the performances of the two different shaping techniques are presented. For $n = 162$, sphere shaping requires 0.85 dB less SNR than CCDM to achieve 10^{-3} frame error probability while it drops to 0.35 dB for $n = 486$. This justify our earlier statement that as n decreases, the difference between required block lengths increases in favor of sphere shaping.

5 Conclusion

Two different shaping techniques are discussed: Constant composition distribution matching and n-sphere shaping. It is shown that both asymptotically achieve the maximum rate and induce MB distribution. Sphere codes perform more efficiently (especially for short block lengths), i.e., approach the maximum rate (MB) faster than

[‖]Note that in this work $H(X) = R + 1$.

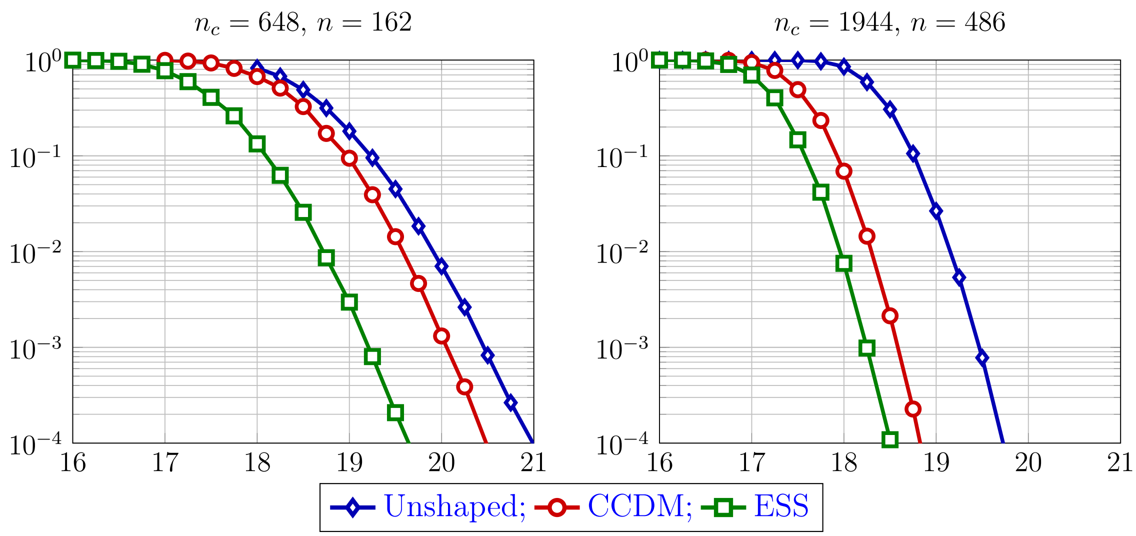

Figure 5: SNR vs. frame error probability for 16-ASK at target rate $R = 2.67$ for $n_c = \{648, 1944\}$ with and without shaping.

constant composition codes. This is shown by investigating the finite-length rate losses. Finally the comparison is justified by presenting Monte Carlo simulations employing LDPC codes.

References

[1] R. F. H. Fischer, *Precoding and Signal Shaping for Digital Transmission*. Wiley-Interscience, 2002.

[2] G. Böcherer and R. Mathar, "Matching Dyadic Distributions to Channels," in *2011 Data Compression Conference*, March 2011, pp. 23–32.

[3] P. Schulte and G. Böcherer, "Constant Composition Distribution Matching," *IEEE Transactions on Information Theory*, vol. 62, no. 1, pp. 430–434, Jan 2016.

[4] G. Böcherer, F. Steiner, and P. Schulte, "Bandwidth Efficient and Rate-Matched Low-Density Parity-Check Coded Modulation," *IEEE Trans. on Commun.*, vol. 63, no. 12, pp. 4651–4665, Dec 2015.

[5] Y. C. Gültekin, W. J. van Houtum, S. Şerbetli, and F. M. J. Willems, "Constellation Shaping for IEEE 802.11," in *2017 IEEE 28th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*, Oct 2017, pp. 1–7.

[6] G. Böcherer and B. C. Geiger, "Optimal Quantization for Distribution Synthesis," *IEEE Transactions on Information Theory*, vol. 62, no. 11, pp. 6162–6172, Nov 2016.

[7] T. V. Ramabadran, "A Coding Scheme for m-out-of-n Codes," *IEEE Transactions on Communications*, vol. 38, no. 8, pp. 1156–1163, Aug 1990.

[8] T. M. Cover and J. A. Thomas, *Elements of Information Theory (Wiley Series in Telecommunications and Signal Processing)*. Wiley-Interscience, 1991.

[9] F. Willems and J. Wuijts, "A Pragmatic Approach to Shaped Coded Modulation," in *IEEE 1st Symp. on Commun. and Veh. Technol. in the Benelux*, 1993.

[10] R. Laroia, N. Farvardin, and S. A. Tretter, "On Optimal Shaping of Multidimensional Constellations," *IEEE Transactions on Information Theory*, vol. 40, no. 4, pp. 1044–1056, Jul 1994.

[11] G. R. Lang and F. M. Longstaff, "A Leech Lattice Modem," *IEEE Journal on Selected Areas in Communications*, vol. 7, no. 6, pp. 968–973, Aug 1989.

[12] G. D. Forney, L. Brown, M. V. Eyuboglu, and J. L. Moran, "The V.34 High Speed Modem Standard," *IEEE Communications Magazine*, vol. 34, no. 12, pp. 28–33, Dec 1996.

[13] T. Prinz, P. Yuan, G. Böcherer, F. Steiner, O. İşcan, R. Böhnke, and W. Xu, "Polar Coded Probabilistic Amplitude Shaping for Short Packets," in *2017 IEEE 18th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, July 2017, pp. 1–5.

[14] P. Schulte and F. Steiner, "Shell Mapping for Distribution Matching," *ArXiv e-prints*, Mar. 2018. [Online]. Available: http://arxiv.org/abs/1803.03614

[15] T. Fehenberger, D. S. Millar, T. Koike-Akino, K. Kojima, and K. Parsons, "Partition-Based Distribution Matching," *ArXiv e-prints*, Jan. 2018. [Online]. Available: https://arxiv.org/abs/1801.08445

[16] "IEEE Standard for Information technology–Telecommunications and Information Exchange Between Systems Local and Metropolitan Area Networks–Specific Requirements - Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications," *IEEE Std 802.11-2016 (Revision of IEEE Std 802.11-2012)*, pp. 1–3534, Dec 2016.