Finite-Support Capacity-Approaching Distributions for AWGN Channels

Derek Xiao*, Linfang Wang*, Dan Song*, Richard D. Wesel*

*University of California, Los Angeles, Los Angeles, CA 90095, USA

Email: derekxiao93@ucla.edu, lfwang@ucla.edu, dansong@ucla.edu , wesel@ucla.edu

***Abstract*—Previously, dynamic-assignment Blahut-Arimoto (DAB) was used to find capacity-achieving probability mass functions (PMFs) for binomial channels and molecular channels. As it turns out, DAB can efficiently identify capacity-achieving PMFs for a wide variety of channels. This paper applies DAB to power-constrained (PC) additive white Gaussian Noise (AWGN) Channels and amplitude-constrained (AC) AWGN Channels.**

This paper modifies DAB to include a power constraint and finds low-cardinality PMFs that approach capacity on PC-AWGN Channels. While a continuous Gaussian PDF is well-known to be capacity-achieving on the PC-AWGN channel, DAB identifies low-cardinality PMFs within 0.01 bits of the mutual information provided by a Gaussian PDF. Recall the results of Ozarow and Wyner requiring a constellation cardinality of $\lceil 2^{C+1} \rceil$ to approach capacity C to within the asymptotic shaping loss of 1.53 dB at high SNR. PMF's found by DAB approach capacity with *essentially no shaping loss* with cardinality less than $2^{C+1.2}$. As expected, DAB's numerical approach identifies PMFs with better mutual information vs. SNR performance than the analytical approaches to finite-support constellations examined by Wu and Verdu.

This paper also uses DAB to find capacity-achieving PMFs with small cardinality support sets for AC-AWGN Channels. The resulting evolution of capacity-achieving PMFs as a function of SNR is consistent with the approximate cardinality transition points of Sharma and Shamai.

I. Introduction

A. Background

Probabilistic amplitude shaping (PAS) [1]–[5] enables practical coded modulation for any symmetric constellation. The constellation points need not be equally spaced or equally likely. PAS awakens interest in low-cardinality probability mass functions (PMFs) that closely approach capacity.

Smith [6] proved that the capacity-achieving distribution for an amplitude-constrained (AC) additive white Gaussian noise (AWGN) channel is unique and has finite-support. Alternative proofs, e.g. [7], [8], followed. Smith also described an algorithm for finding the AC-AWGN capacity-achieving distributions for any amplitude constraint for a fixed noise power. Bounds on capacity-achieving input cardinality were identified by Dytso *et al.* [8] and Yagli *et al.* [9]. Sharma and Shamai [10] derived analytical conditions for SNRs at which the input cardinality increases, including approximate SNRs for transitions from binary and ternary optimal signaling to the next higher cardinality. Finally, various upper and lower bounds on the AC-AWGN capacity were identified [11], [12].

For the power-constrained AWGN (PC-AWGN) channel, Ungerboeck showed empirically [13] and Ozarow and Wyner proved analytically [14] that one-dimensional constellations with 2^{C+1} points could achieve rates within the shaping loss (around 0.25 bits or 1.53 dB for high SNRs) of the channel capacity. The combination of Ungerboeck-cardinality equidistant constellations with probabilistic shaping then resolves most of this loss (0.1 dB or less) for the PC-AWGN channel [15]. Wu and Verdu [16] prove that for any fixed SNR, cardinality-m PMFs defined by Gauss quadratures achieve information rates that converge to the PC-AWGN channel capacity exponentially fast as m increases.

An algorithmic approach with the cutting plane algorithm is proven by Huang and Meyn [17] to converge to the discrete capacity achieving distribution for a variety of channels including the PC-AWGN Channel. Fragouli et al. [18] shows that non-uniform point placement can improve performance for equally likely constellation points. Mathar and Alirezaei [19] argue that discrete distributions can always achieve or closely approach the PC-AWGN capacity.

B. Contributions and Organization

The dynamic assignment Blahut-Arimoto (DAB) algorithm is an iterative algorithm originally used to find the capacity and capacity-achieving finite-support PMFs for the binomial and molecular channels [20], [21].

As an initial contribution, this paper modifies the DAB algorithm to accommodate continuous outputs and use it to identify capacity-achieving finite-support PMFs for the AC-AWGN channel. Our results match Smith [6] and explore how the capacity achieving PMF evolves with a fixed amplitude constraint as noise power varies.

As its main contribution, this paper incorporates a power-constraint into DAB[1], and uses it to find cardinality-constrained input PMFs that can closely approach the capacity of the PC-AWGN channel, C = $\frac{1}{2}\log_2(1+\mathrm{SNR})$. Despite the well-known result that the PC-AWGN capacity is achieved by a continuous (Gaussian) PDF, our results demonstrate that a discrete input PMF can always approach the capacity within less than 0.01 bits as long as $\log_2$ of the input cardinality

This research is supported in part by National Science Foundation (NSF) grant CCF-1911166. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the NSF.

[1] Code for both versions of DAB can be found at https://github.com/UCLA-Communications-Systems-Lab/DAB-for-AWGN-channels

is 1.2 bits above the PC-AWGN capacity, or, alternatively, the input entropy is 0.9 bits above the PC-AWGN capacity. This is a significant reduction in input cardinality compared to analytical distributions from [16], for which the required excess log input cardinality above capacity to achieve a 0.01 shaping loss apparently increases without bound for higher SNR.

This paper is organized as follows: Sec. II introduces DAB for AC-AWGN channels and obtains finite-support PMFs that closely approach the AC-AWGN capacity. Sec. III adapts the DAB approach to PC-AWGN channels by incorporating a power constraint, and provides minimum-cardinality PMFs that closely approach the PC-AWGN capacity for a wide range of SNRs. Sec. IV concludes the paper.

II. AWGN Channel with an Amplitude Constraint

This section presents a modified version of the DAB algorithm introduced in [21] that can be applied to the AC-AWGN channel. For this channel, input X with support $[-1, 1]$ is corrupted by AWGN to produce the continuous output Y with PDF $f(y)$. The input and output alphabets are continuous, but a unique finite-support capacity-achieving PMF for X is known to exist [6]. The original DAB algorithm assumed a discrete output alphabet, so modifications are required to compute some information theoretic quantities with continuous output alphabets.

Alg. 1 summarizes the basic steps of DAB for the AC-AWGN channel. The general structure of the DAB algorithm throughout this paper comprises the following steps in each iteration: using Blahut-Arimoto to optimize the allocation of probability to mass points, checking for convergence, adding a mass point if necessary, and optimizing mass point locations.

A. DAB for the AC-AWGN channel

For the AC-AWGN channel, define $\gamma = 1/N$ where N is the variance of the AWGN. For the AC-AWGN channel with a specified γ, DAB starts by initializing the number of mass points $|\mathcal{X}^{(1)}|$ and their locations $\mathcal{X}^{(1)}$ to optimal values previously obtained by DAB for the $\gamma - \Delta\gamma$ case. For $\Delta\gamma$ small enough, the capacity-achieving cardinality $|\mathcal{X}^{*}|$ is either $|\mathcal{X}^{(1)}|$ or $|\mathcal{X}^{(1)}| + 1$, as seen in Fig. 1.

The details of the four major steps of Alg. 1 are as follows:

1) Optimizing Probability Assignment $\mathcal{P}^{(k)}$: The Blahut-Arimoto Algorithm [22] finds the probability assignment $\mathcal{P}^{(k)}$ over the current support set $\mathcal{X}^{(k)}$ that maximizes mutual information (MI). The MI $I^{(k)}$ is a lower bound on the AC-AWGN capacity.

2) Checking for Convergence: At the heart of DAB is Csiszar's Min-Max Capacity Theorem [23], which states:

$$C = \min_{f_Y(\cdot)\in\{F_Y\}} \max_x D\big(f_{Y|X}(\,\cdot\,|x)\,\|\, f_Y(\,\cdot\,)\big)\,, \tag{1}$$

where $\{F_Y\}$ is the set of distributions on Y induced by a valid input probability distribution. An upper bound on capacity follows from (1): For any valid output distribution on $\mathcal{Y}$,

$$C \le D^{(k)}_{\text{max}} = \max_x D\left(f_{Y|X}(\,\cdot\,|x)\,\|\, f_Y(\,\cdot\,)\right)\,. \tag{2}$$

Algorithm 1 AC-AWGN DAB

Initialization: Select the tolerance ϵ, which is the maximum acceptable distance from capacity. For our results $\epsilon = 10^{-5}$. For a specified value of $\gamma = 1/N$, $\mathcal{X}^{(1)}$ and $|\mathcal{X}^{(1)}|$ are set to the values found previously by DAB for the $\gamma - \Delta\gamma$ case. The mass point locations are arranged in increasing order in the vector $\mathcal{X}^{(1)} = \begin{bmatrix} x_1 & x_2 & \ldots & x_{|\mathcal{X}^{(1)}|}\end{bmatrix}$.

Iterations: Determination of the optimal $\mathcal{X}^{*}$, $\mathcal{P}^{*}$, and the capacity C (within ϵ bits) proceeds as follows:

1) *Optimize probability assignments $\mathcal{P}^{(k)}$:* Given $\mathcal{X}^{(k)}$, use Blahut-Arimoto to compute the MI-maximizing PMF $\mathcal{P}^{(k)}$ and corresponding MI $I^{(k)}$, a lower bound on C.
2) *Check for Convergence:*
   a) Use the distribution $f(y)$ induced by $\mathcal{X}^{(k)}$ and $\mathcal{P}^{(k)}$ to compute the capacity upper bound
   $$D^{(k)}_{\text{max}} = \max_{x\in\mathcal{X}} D\big(f_{Y|X}(\,\cdot\,|x)\,\|\, f_Y(\,\cdot\,)\big)$$
   and $x^{(k)}_{\text{max}}$ which maximizes the upper bound
   $$x^{(k)}_{\text{max}} = \arg\max_x D\big(f_{Y|X}(\,\cdot\,|x)\,\|\, f_Y(\,\cdot\,)\big)\,.$$
   b) If $D^{(k)}_{\text{max}} - I^{(k)} < \epsilon$ terminate with $\mathcal{X}^{*} = \mathcal{X}^{(k)}$, $\mathcal{P}^{*} = \mathcal{P}^{(k)}$, and $C = I^{(k)}$. Otherwise continue.
3) *Add a mass point if needed:* If there is a mass point in $(0, x^{(k)}_{\text{max}})$ then $|\mathcal{X}^{(k+1)}| = |\mathcal{X}^{(k)}|$. Otherwise, $|\mathcal{X}^{(k+1)}| = |\mathcal{X}^{(k)}|+1$, and $\mathcal{X}^{(k+1)}$ is updated to include the additional location.
4) *Optimize mass point locations $\mathcal{X}$:*
   a) Determine direction vector $\tilde{\mathcal{D}}^{k}$ to adjust $\mathcal{X}$.
   b) Compute
   $$\mathcal{X}^{(k+1)} = \mathcal{X}^{(k)} + \lambda^{*}\tilde{\mathcal{D}}^{k}\,,$$
   where
   $$\lambda^{*} = \arg\max_{\lambda} I\left(\mathcal{X}^{(k)} + \lambda\tilde{\mathcal{D}}^{k}, \mathcal{P}^{(k)}\right)\,,$$
   and $I(\mathcal{X}, \mathcal{P})$ is the mutual information that results from an input PMF characterized by $\mathcal{X}$ and $\mathcal{P}$.
5) $\gamma = \gamma + \Delta\gamma$
6) Go to 1.

In each DAB iteration, Eq. (2) is used to update $D^{(k)}_{\text{max}}$. The algorithm terminates if difference of the upper and lower bounds $D^{(k)}_{\text{max}} - I^{(k)}$ is less than the tolerance ϵ.

3) Adding a Mass Point if Needed: Motivated by the point optimization in step 4, an additional mass point is added if no mass point lies in the open interval $(0, x^{(k)}_{\text{max}})$ where

$$x^{(k)}_{\text{max}} = \arg\max_x D\big(f_{Y|X}(\,\cdot\,|x)\,\|\, f_Y(\,\cdot\,)\big)\,. \tag{3}$$

If $|\mathcal{X}^{(k)}|$ is even, a new point is added at 0. If $|\mathcal{X}^{(k)}|$ is odd, a new point is added by splitting the point at 0 into two points that move away from 0 in the line search of step 4b.

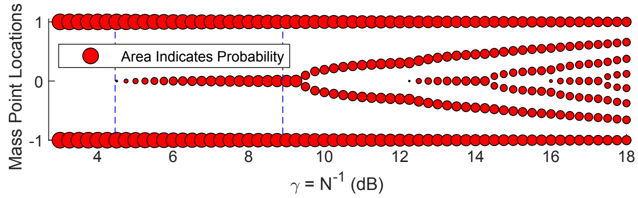


Fig. 1: Capacity-achieving input PMFs of the AC-AWGN channel. The approximate theoretical input cardinality transition points [10] are shown with vertical blue dotted lines.

4) Optimizing Point Locations $\mathcal{X}$: First, a direction $\tilde{\mathcal{D}}^k$ is selected along which $\mathcal{X}^k$ will be varied in step 4b to increase the mutual information $I(\mathcal{X}, \mathcal{P})$, computed through MATLAB's integral() function. Several approaches to selecting $\tilde{\mathcal{D}}^k$ are presented in [21]. The approach we adopt selects $\tilde{\mathcal{D}}^k = e_j - e_{|\mathcal{X}^{(k)}|+1-j}$ to move a symmetric pair of mass points, where e_j is the j-th standard basis vector, since the AC-AWGN capacity-achieving distribution is symmetric [6].

We choose j so that x_j is the point in the interval (0, $x_{\max}^{(k)}$) closest to $x_{\max}^{(k)}$ as defined by (3) to lower the capacity upper bound. A line search routine (such as fminbnd in MATLAB) finds the λ that maximizes $I(\mathcal{X}, \mathcal{P})$ in step 4b of Alg. 1.

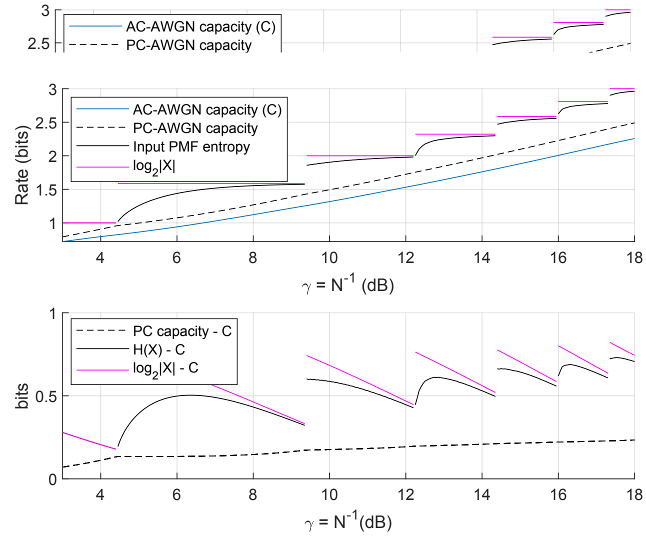
B. AC-AWGN Capacity-Achieving PMFs

This section presents capacity-approaching PMFs for the AC-AWGN channel, compares the AC-AWGN channel capacity and the PC-AWGN channel capacity, and explores how the input cardinality for capacity-achieving PMFs for the AC-AWGN channel evolves over a wide range of SNRs.

The DAB-optimized input PMFs approach capacity within a specified small tolerance ϵ. We use $\epsilon = 10^{-5}$ bits. Fig. 1 shows the resulting PMFs plotted against $\gamma = \frac{1}{N}$, for noise power N. Input cardinality transitions in Fig. 1 where binary and ternary signaling become sub-optimal occur at $\gamma = 4.44$ dB and 9.28 dB respectively. The approximate analysis in [10] place the transitions at $\gamma = 4.46$ and 8.90 dB.

Fig. 2a shows the AC-AWGN capacity and the PC-AWGN capacity for the SNR resulting from the input PMF found by DAB. This illustrates the rate loss resulting from the amplitude constraint, which is small at low SNR. Fig. 2a also shows $\log_2 |\mathcal{X}|$ and the input entropy $H(x)$. Both the cardinality and entropy of X can be viewed as resources, and these two resource requirements grow in a predictable way.

Fig. 2b shows the same information as Fig. 2a but with the AC-AWGN capacity subtracted. The $\log_2 |\mathcal{X}| - C$ curve for the AC-AWGN channel exhibits a maximum "excess cardinality" above the AC-AWGN capacity of around 0.8 bits. Large discontinuous jumps in input entropy occur when an odd number of mass points transitions to an even number. The difference between the AC-AWGN capacity and PC-AWGN capacity appears to asymptotically approach 0.25 bits, which is the shaping loss of equilattice constellations, where points are equally spaced and equally likely.



(b) Each of the curves in Fig. 2a, with C subtracted.

Fig. 2: Cardinality and rate for the AC-AWGN capacity-achieving input PMFs

III. AWGN Channel with a Power Constraint

In Subsection III-A, we present a modified DAB algorithm for finding minimum cardinality capacity-approaching input PMFs for the PC-AWGN channel. In Subsection III-B, we demonstrate that a finite-support input PMF approaches the PC-AWGN capacity within less than 0.01 bits as long as $\log_2 |\mathcal{X}|$ is 1.2 bits above the PC-AWGN capacity. We also quantify the improvement in shaping loss compared to the four types of analytical capacity-approaching input PMFs proposed by Wu and Verdu in [16].

A. DAB for the PC-AWGN channel

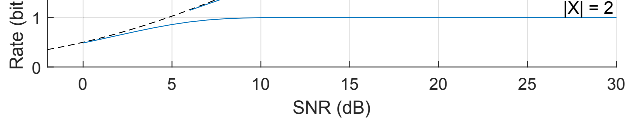
The modifications to each step of Alg. 1 are motivated by the requirement that the input PMF satisfies a power constraint E. Each step in the algorithm, therefore, must search only among PMFs that have an average power of E (or less).

1) Optimizing Probability Assignment: When finding the optimal probability assignment $\mathcal{P}^{(k)}$, we use the version of the Blahut-Arimoto algorithm [22] that finds the capacity with a power constraint parameterized by s, for fixed point locations.

While E decreases monotonically with s, the power of the probability assignment resulting from choosing a particular s is unknown before the Blahut-Arimoto algorithm terminates, so we use the secant method to find the value of s that produces an optimal probability assignment with the required power constraint within a specified tolerance of 10^{-8}.

2) Checking for Convergence: We monitored convergence by measuring the rate improvement per iteration, where the algorithm terminates if the mutual information increase per iteration drops below a predefined tolerance of 10^{-5}.

3) Adding a Mass Point if Needed: For the PC-AWGN channel, we did not add mass points within the DAB algorithm. Instead, for each fixed cardinality, DAB finds the



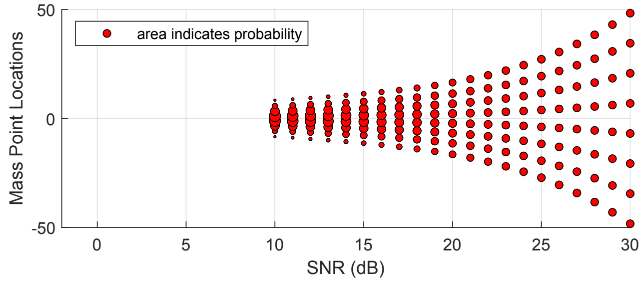


Fig. 3: DAB-optimized input PMFs for $|\mathcal{X}| = 8$.

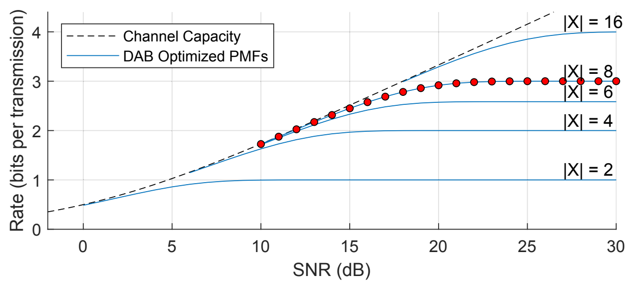


Fig. 4: Information rates for the DAB optimized input PMFs for $|\mathcal{X}| = 2, 4, 6, 8,$ and 16 for the PC-AWGN channel, which approach channel capacity to a practically negligible shaping loss at low enough SNR.

optimal input PMFs in the interval of interesting SNRs. For each cardinality, this produces an evolution of input PMFs, such as the one in Fig. 3 with the corresponding mutual informations shown as red dots in Fig. 4. This is then repeated for the input cardinalities of interest, which could be limited to powers of two or could be all integer values in light of PAS coded modulation. The lowest cardinality input PMF which still achieves a 0.01 bit gap to capacity is selected as the "capacity-approaching input PMF" according to DAB.

4) Optimizing Point Locations: Moving a symmetric mass point pair without adjusting probability violates the power constraint or leaves power unused. To meet the power constraint with equality, let p_i and x_i be the i^{th} element of $\mathcal{P}^{(k)}$ and $\mathcal{X}^{(k)} + \lambda\tilde{\mathcal{D}}^k$ respectively. The symmetric mass point pair with index ℓ and $r = |\mathcal{X}^{(k)}| - \ell + 1$, must satisfy

$$\alpha_s\left(p_\ell + p_r\right) + \alpha_o \sum_{i \notin [\ell, r]} p_i + \sum_{\ell < i < r} p_i = 1 \quad (4)$$

$$\alpha_s\left(p_\ell x_\ell^2 + p_r x_r^2\right) + \alpha_o \sum_{i \notin [\ell, r]} p_i x_i^2 + \sum_{\ell < i < r} p_i x_i^2 = E \quad (5)$$

where factors α_s and α_o scale the symmetric and outer probabilities. We use the solution to define the new probability assignment $\mathcal{P}_\lambda^{(k)}$. The process for outermost points is analogous, transferring probability with inner points. The optimized point pair changes in a round robin fashion, but only one point pair is optimized in each iteration.

B. Minimum Cardinality Capacity-Approaching PMFs

Fig. 3 shows DAB-optimized input PMFs for $|\mathcal{X}| = 8$ with the corresponding mutual information shown as red dots in

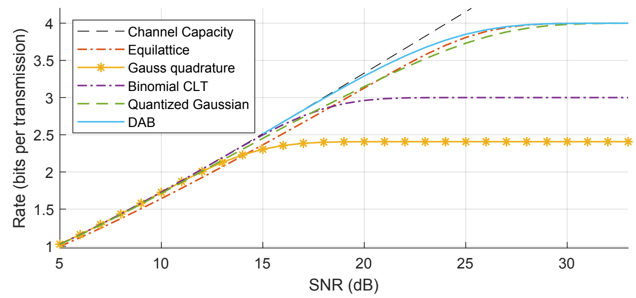


Fig. 5: Shaping loss of DAB optimized input PMFs with cardinality $|\mathcal{X}| = 16$ is lower than the shaping loss of the four analytical input PMFs from Wu and Verdu [16].

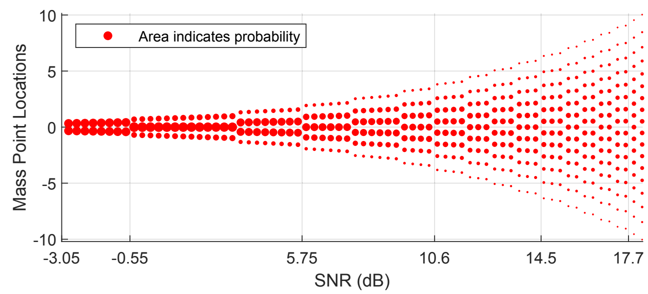


Fig. 6: Minimum cardinality capacity-approaching input PMFs (less than 0.01 bits of shaping loss from capacity).

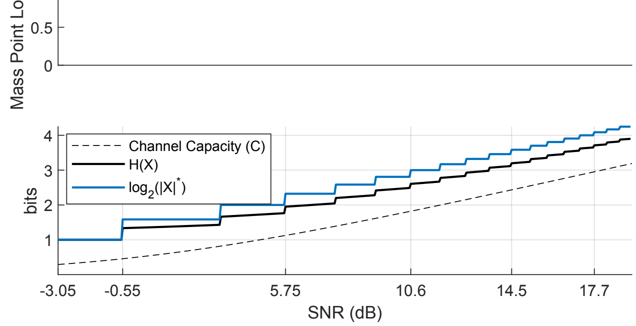
Fig. 4: The optimal PMFs for a fixed-cardinality tend towards an equilattice as SNR increases.

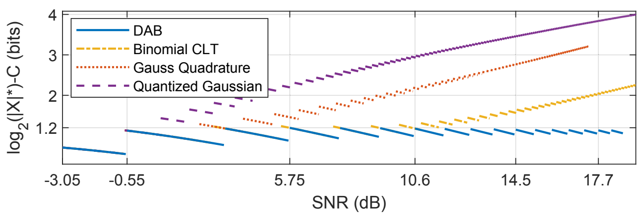
Fig. 4 shows the mutual information values achieved by DAB PMFs for $|\mathcal{X}| = 2, 4, 6, 8, 16$. The loss from capacity (shaping loss) rapidly approaches 0 bits as SNR decreases.

Focusing on $|\mathcal{X}| = 16$ as an example, Fig. 5 shows how the DAB optimized input PMFs compare to the four analytical input PMFs from Wu and Verdu [16] over a large range of SNRs. The theoretical proofs of optimality in [16] apply in the limit of low-SNR(with Gauss Quadrature) and high-SNR(with equilattice), and those input PMFs do well at there respective optimal extremes. However, DAB expectedly gives superior results across the full SNR range. Unlike the PMFs of Wu and Verdu, DAB achieves a mutual information very close to capacity for the practically important operating point where $\log|\mathcal{X}|$ is about a bit above the capacity. We see similar behavior comparing DAB to Wu and Verdu PMFs for $|\mathcal{X}| = 2, 4, 6, 8, 16,$ and 32.

Fig. 6 shows the input PMFs resulting from using DAB to find the minimum cardinality required to achieve less than a 0.01 bit shaping loss for SNRs from -3 to 19 dB and all input cardinalities from 2 to 19.

Fig. 7 explores the requirements in terms of cardinality and input entropy of the capacity-approaching DAB PMFs shown in. Fig. 6, and compares this to the PMFs of Wu and Verdu. Fig. 7a plots $\log_2(|\mathcal{X}^*|)$ and input entropy for the DAB PMFs, and both tend toward a constant offset above





(b) $\log_2 |\mathcal{X}^*| - C$ for optimal DAB distributions, and three analytical distributions from Wu and Verdu [16].

Fig. 7: Capacity approaching PMFs and a bound on log input cardinality to achieve a shaping loss of 0.01 bits.

capacity. Fig. 7b shows the offset $\log_2(|\mathcal{X}^*|) - C$ for input PMFs that approach capacity within 0.01 bits, showing the DAB PMFs and only three of the four analytical distributions from [16] since the equilattice cannot approach capacity to within 0.01 bits. The cardinality offset required to achieve a 0.01 bit shaping loss with DAB is upper bounded by 1.2 bits, while the cardinality required to achieve this shaping loss increases linearly for the three analytical distributions. So, significant reductions in constellation complexity can be obtained by using an iteratively optimized input distribution instead of the analytically inspired input distributions.

IV. Conclusion

As its main contribution, this paper modifies the DAB algorithm and uses it to find minimum-cardinality capacity-approaching input PMFs for the PC-AWGN channel. Huang and Meyn [17] showed examples where small cardidality constellations can do well, but DAB finds the smallest cardinality finite-support PMF within a specified negligibly small distance (0.01 bits) from the PC-AWGN channel capacity. These input PMFs have a cardinality of less than $2^{C+1.2}$ and demonstrate a significant reduction in input cardinality compared to analytical distributions from [16] for high SNRs.

This paper also applies the DAB algorithm to AC-AWGN channels where the capacity-achieving PMF is proven to be finite-support and unique [6]. DAB provides a complete illustration of how the capacity achieving PMFs evolve as a function of the ratio γ of squared amplitude constraint to noise variance. We observe that the optimal input cardinality $|\mathcal{X}|$ is close to the highest $|\mathcal{X}|$ for which $\log_2 |\mathcal{X}| - C \leq 0.8$.

References

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

[2] G. Böcherer, “Achievable Rates for Probabilistic Shaping,” *arXiv:1707.01134v5 [cs, math]*, May 2018.

[3] P. Schulte, F. Steiner, and G. Böcherer, “Four Dimensional Probabilistic Shaping for Fiber-Optic Communication,” in *Adv. Photon. 2017 (IPR, NOMA, Sensors, Netw., SPPCom, PS) (2017), Paper SpM2F.5*. Optical Society of America, Jul. 2017, p. SpM2F.5.

[4] F. Buchali, F. Steiner, G. Böcherer, L. Schmalen, P. Schulte, and W. Idler, “Rate Adaptation and Reach Increase by Probabilistically Shaped 64-QAM: An Experimental Demonstration,” *J. Lightw. Technol.*, vol. 34, no. 7, pp. 1599–1609, Apr. 2016.

[5] G. Böcherer, P. Schulte, and F. Steiner, “Probabilistic Shaping and Forward Error Correction for Fiber-Optic Communication Systems,” *Journal of Lightwave Technology*, vol. 37, no. 2, pp. 230–244, Jan. 2019. [Online]. Available: https://ieeexplore.ieee.org/document/8627924/

[6] J. G. Smith, “The information capacity of amplitude- and variance-constrained scalar gaussian channels,” *Inf. Control.*, vol. 18, pp. 203–219, 1971.

[7] T. Chan, S. Hranilovic, and F. Kschischang, “Capacity-achieving probability measure for conditionally gaussian channels with bounded inputs,” *IEEE Trans. Inf. Theory*, vol. 51, no. 6, pp. 2073–2088, June 2005.

[8] A. Dytso, S. Yagli, H. V. Poor, and S. S. Shitz, “The capacity achieving distribution for the amplitude constrained additive gaussian channel: An upper bound on the number of mass points,” *IEEE Trans. Inf. Theory*, vol. 66, no. 4, pp. 2006–2022, Apr. 2020.

[9] S. Yagli, A. Dytso, and H. V. Poor, “Estimation of bounded normal mean: An alternative proof for the discreteness of the least favorable prior,” in *2019 IEEE Inf. Theory Workshop (ITW)*. IEEE, Aug. 2019.

[10] N. Sharma and S. Shamai, “Characterizing the discrete capacity achieving distribution with peak power constraint at the transition points,” *2008 International Symp. on Inf. Theory and Its Apps.*, Dec. 2008.

[11] A. Thangaraj, G. Kramer, and G. Böcherer, “Capacity Bounds for Discrete-Time, Amplitude-Constrained, Additive White Gaussian Noise Channels,” *IEEE Trans. Inf. Theory*, vol. 63, no. 7, pp. 4172–4182, Jul. 2017.

[12] A. McKellips, “Simple tight bounds on capacity for the peak-limited discrete-time channel,” in *Int. Symp. Inf. Theory, 2004. ISIT 2004. Proceedings*. IEEE, 2004.

[13] G. Ungerboeck, “Channel coding with multilevel/phase signals,” *IEEE Trans. Inf. Theory*, vol. 28, no. 1, pp. 55–67, Jan. 1982.

[14] L. Ozarow and A. Wyner, “On the capacity of the gaussian channel with a finite number of input levels,” *IEEE Trans. Inf. Theory*, vol. 36, no. 6, pp. 1426–1428, 1990.

[15] R. Modonesi, M. Dalai, P. Migliorati, and R. Leonardi, “A Note on Probabilistic and Geometric Shaping for the AWGN Channel,” *IEEE Communications Letters*, vol. 24, no. 10, pp. 2119–2122, Oct. 2020, conference Name: IEEE Communications Letters.

[16] Y. Wu and S. Verdu, “The impact of constellation cardinality on gaussian channel capacity,” in *2010 48th Annual Allerton Conference on Comm., Control, and Computing (Allerton)*. IEEE, Sep. 2010.

[17] J. Huang and S. Meyn, “Characterization and computation of optimal distributions for channel coding,” *IEEE Trans. Inf. Theory*, vol. 51, no. 7, pp. 2336–2351, July 2005.

[18] C. Fragouli, R. D. Wesel, D. Sommer, and G. P. Fettweis, “Turbo codes with non-uniform constellations,” in *ICC 2001. IEEE International Conference on Communications. Conference Record (Cat. No.01CH37240)*, vol. 1, 2001, pp. 70–73 vol.1.

[19] G. Alirezaei and R. Mather, “There always is a discrete capacity-achieving distribution,” *2020 Inf. Theory and Apps. Workshop*, 2020.

[20] Richard D. Wesel, Emily E. Wesel, Lieven Vandenberghe, Christos Komninakis, and Muriel Medard, “Efficient Binomial Channel Capacity Computation with an Application to Molecular Communication,” in *Proc. Inf. Theory and Apps (ITA) Workshop*, La Jolla (CA), Feb. 2018.

[21] N. Farsad, W. Chuang, A. Goldsmith, C. Komninakis, M. Médard, C. Rose, L. Vandenberghe, E. E. Wesel, and R. D. Wesel, “Capacities and optimal input distributions for particle-intensity channels,” 2020, arXiv:2005.10682 [cs.IT].

[22] R. Blahut, “Computation of channel capacity and rate-distortion functions,” *IEEE Trans. Inf. Theory*, vol. 18, no. 4, pp. 460–473, Jul. 1972.

[23] I. Csisár and J. Körner, *Information Theory: Coding Theorems for Discrete Memoryless Systems*, ser. Probability and Mathematical Statistics, Z. Birnbaum and E. Lukacs, Eds. New York - San Francisco - London: Academic Press, 1981. See Theorem 3.4.