Greedy-Merge Degrading has Optimal Power-Law

Assaf Kartowsky and Ido Tal
Department of Electrical Engineering
Technion - Haifa 32000, Israel
E-mail: {kartov@campus, idotal@ee}.technion.ac.il

Abstract—To be considered for the 2017 IEEE Jack Keil Wolf ISIT Student Paper Award.

Consider a channel with a given input distribution. Our aim is to degrade it to a channel with at most L output letters. One such degradation method is the so called "greedy-merge" algorithm. We derive an upper bound on the reduction in mutual information between input and output. For fixed input alphabet size and variable L, the upper bound is within a constant factor of an algorithm-independent lower bound. Thus, we establish that greedy-merge is optimal in the power-law sense.

I. INTRODUCTION

In myriad digital processing contexts, quantization is used to map a large alphabet to a smaller one. For example, quantizers are an essential building block in receiver design, used to keep the complexity and resource consumption manageable. The quantizer used has a direct influence on the attainable code rate.

Another recent application is related to polar codes [1]. Polar code construction is equivalent to evaluating the misdecoding probability of each channel in a set of synthetic channels. This evaluation cannot be carried out naively, since the output alphabet size of a synthetic channel is intractably large. One approach to circumvent this difficulty is to degrade the evaluated synthetic channel to a channel with manageable output alphabet size [2][3][4][5][6].

Given a design parameter L, we degrade an initial channel to a new one with output alphabet size at most L. We assume that the input distribution is specified, and note that this degradation reduces the mutual information between the channel input and output, respectively. In both examples above, this reduction is roughly the loss in code rate due to quantization. We denote the smallest reduction possible by ΔI^*.

Let $|\mathcal{X}|$ denote the channel input alphabet size, and treat it as a fixed quantity. We show that for any input distribution and any initial channel, $\Delta I^* = O(L^{-2/(|\mathcal{X}|-1)})$. Moreover, this bound is attained efficiently, by the greedy-merge algorithm [2][5]. This bound is tighter than the bounds derived in [3], [4], [5] and [6]. In fact, up to constant multipliers (dependent on $|\mathcal{X}|$), this bound is the tightest possible. Namely, [7] proves the existence of an input distribution and a sequence of channels for which $\Delta I^* = \Omega(L^{-2/(|\mathcal{X}|-1)})$. Both bounds have $-2/(|\mathcal{X}| - 1)$ as the power of L, the same power-law.

II. PRELIMINARIES

We are given an input distribution and a discrete memoryless channel (DMC) $W : \mathcal{X} \to \mathcal{Y}$. Both $|\mathcal{X}|$ and $|\mathcal{Y}|$ are assumed finite. Let X and Y denote the random variables that correspond to the channel input and output. Denote the corresponding distributions P_X and P_Y. Let $W(y|x) \triangleq \mathbb{P}\{Y = y|X = x\}$. For brevity, let $\pi(x) \triangleq \mathbb{P}\{X = x\} = P_X(x)$. Assuming further that $\mathcal{X}$ and $\mathcal{Y}$ are disjoint, we abuse notation and denote $\mathbb{P}\{X = x|Y = y\}$ and $\mathbb{P}\{Y = y\}$ as $W(x|y)$ and $\pi(y)$, respectively. Without loss of generality, $\pi(x) > 0$ and $\pi(y) > 0$. We *do not* assume that W is symmetric.

The mutual information between channel input and output is

$$I(W) \triangleq I(X;Y) = \sum_{x \in \mathcal{X}} \eta(\pi(x)) - \sum_{\substack{x \in \mathcal{X}, \\ y \in \mathcal{Y}}} \pi(y)\eta(W(x|y)) \,,$$

where $\eta(p) \triangleq -p \log p$ for $p > 0$, zero for $p = 0$, and the logarithm is taken in the natural basis. We note that the input distribution does not necessarily have to be the one that achieves the channel capacity.

We now define the relation of degradedness between channels. A channel $Q : \mathcal{X} \to \mathcal{Z}$ is said to be (stochastically) *degraded* with respect to a channel $W : \mathcal{X} \to \mathcal{Y}$, and we write $Q \preccurlyeq W$, if there exists a channel $\Phi : \mathcal{Y} \to \mathcal{Z}$ such that

$$Q(z|x) = \sum_{y \in \mathcal{Y}} W(y|x)\Phi(z|y) \,, \tag{1}$$

for all $x \in \mathcal{X}$ and $z \in \mathcal{Z}$. Note that as a result of the data processing theorem, $Q \preccurlyeq W$ implies $\Delta I \triangleq I(W) - I(Q) \geq 0$.

Although mentioned before, let us properly define the *optimal degrading loss* for a given pair (W, P_X) as

$$\Delta I^* \triangleq \min_{\substack{Q, \Phi : Q \preccurlyeq W, \\ |Q| \leq L}} I(W) - I(Q) \,, \tag{2}$$

where $|Q|$ denotes the output alphabet size of the channel Q. The optimizer Q is the degraded channel that is "closest" to W in the sense of mutual information, yet has at most L output letters.

III. UPPER BOUND ON OPTIMAL DEGRADING LOSS

A. *Main result*

Our main result is an upper bound on ΔI^*, in terms of $|\mathcal{X}|$ and L. This upper bound will follow from analyzing a sub-optimal[1] degrading algorithm, called "greedy-merge". In each iteration of greedy-merge, we merge the two output letters $y_a, y_b \in \mathcal{Y}$ that results in the smallest decrease of mutual

---

[1]For the binary-input case, optimal degrading can be realized through dynamic programming [8]. For the non-binary case, we do not know of an efficient realization of optimal degrading.

information between input and output, denoted ΔI. Namely, the intermediate channel Φ maps y_a and y_b to a new symbol, while all other symbols are unchanged by Φ. This is repeated $|\mathcal{Y}| - L$ times, to yield an output alphabet size of L. By upper bounding the ΔI of each iteration we obtain an upper bound on ΔI^*. A key result is the following theorem, stating that there exists a pair of output letters whose merger yields a "small" ΔI.

Theorem 1. *Let a DMC $W : \mathcal{X} \to \mathcal{Y}$ satisfy $|\mathcal{Y}| > 2|\mathcal{X}|$. There exists a pair $y_a, y_b \in \mathcal{Y}$ whose merger results in a channel Q satisfying $\Delta I = O\left(|\mathcal{Y}|^{-\frac{|\mathcal{X}|+1}{|\mathcal{X}|-1}}\right)$. In particular,*

$$\Delta I \leq \mu(|\mathcal{X}|) \cdot |\mathcal{Y}|^{-\frac{|\mathcal{X}|+1}{|\mathcal{X}|-1}} , \tag{3}$$

where,

$$\mu(|\mathcal{X}|) \triangleq \frac{\pi |\mathcal{X}|}{\left(\sqrt{1 + \frac{1}{2(|\mathcal{X}|-1)}} - 1\right)^2} \left(\frac{2|\mathcal{X}|}{\Gamma\left(1 + \frac{|\mathcal{X}|-1}{2}\right)}\right)^{\frac{2}{|\mathcal{X}|-1}} ,$$

and $\Gamma(\cdot)$ is the Gamma function.

Recall that Theorem 1 is referring to the merger of a single pair of output letters. The following corollary is our main result, and is basically an iterative utilization of Theorem 1.

Corollary 2. *Let a DMC $W : \mathcal{X} \to \mathcal{Y}$ satisfy $|\mathcal{Y}| > 2|\mathcal{X}|$ and let $L \geq 2|\mathcal{X}|$. Then,*

$$\Delta I^* = \min_{\substack{Q,\Phi: Q \preccurlyeq W, \\ |Q| \leq L}} I(W) - I(Q) = O\left(L^{-\frac{2}{|\mathcal{X}|-1}}\right) .$$

In particular, $\Delta I^* \leq \nu(|\mathcal{X}|) \cdot L^{-\frac{2}{|\mathcal{X}|-1}}$, where $\nu(|\mathcal{X}|) \triangleq \frac{|\mathcal{X}|-1}{2}\mu(|\mathcal{X}|)$, and $\mu(\cdot)$ was defined in Theorem 1. This bound is attained by greedy-merge, and is tight in the power-law sense.

Proof. If $L \geq |\mathcal{Y}|$, then obviously $\Delta I^* = 0$ which is not the interesting case. If $2|\mathcal{X}| \leq L < |\mathcal{Y}|$, then applying Theorem 1 repeatedly $|\mathcal{Y}| - L$ times yields

$$\begin{aligned}
\Delta I^* &\leq \sum_{\ell=L+1}^{|\mathcal{Y}|} \mu(|\mathcal{X}|) \cdot \ell^{-\frac{|\mathcal{X}|+1}{|\mathcal{X}|-1}} \\
&\leq \mu(|\mathcal{X}|) \int_L^{|\mathcal{Y}|} \ell^{-\frac{|\mathcal{X}|+1}{|\mathcal{X}|-1}} \, \mathrm{d}\ell \\
&\leq \nu(|\mathcal{X}|) \cdot L^{-\frac{2}{|\mathcal{X}|-1}} ,
\end{aligned}$$

by the monotonicity of $\ell^{-(|\mathcal{X}|+1)/(|\mathcal{X}|-1)}$. The bound is tight in the power-law sense, by [7, Theorem 2]. ∎

Note that for large values of $|\mathcal{X}|$, the Stirling approximation along with some other first order approximations can be applied to simplify $\nu(|\mathcal{X}|)$ to $\nu(|\mathcal{X}|) \approx 16\pi e |\mathcal{X}|^3$.

B. An alternative "distance" function

We return now to the proof of Theorem 1, and begin by addressing the merger of a pair of output letters $y_a, y_b \in \mathcal{Y}$. Since this pair of letters is merged into a new letter y_{ab} (we assume $y_{ab} \notin \mathcal{X} \cup \mathcal{Y}$), the new output alphabet of Q is $\mathcal{Z} = \mathcal{Y} \setminus \{y_a, y_b\} \cup \{y_{ab}\}$. The channel $Q : \mathcal{X} \to \mathcal{Z}$ then satisfies $Q(y_{ab}|x) = W(y_a|x) + W(y_b|x)$, whereas for all $y \in \mathcal{Z} \cap \mathcal{Y}$ we have $Q(y|x) = W(y|x)$. Using the shorthand

$$\pi_{ab} = \pi(y_{ab}) , \quad \pi_a = \pi(y_a) , \quad \pi_b = \pi(y_b) ,$$

one gets that $\pi_{ab} = \pi_a + \pi_b$. Let us denote by $\boldsymbol{\alpha} = (\alpha_x)_{x \in \mathcal{X}}$, $\boldsymbol{\beta} = (\beta_x)_{x \in \mathcal{X}}$ and $\boldsymbol{\gamma} = (\gamma_x)_{x \in \mathcal{X}}$ the vectors corresponding to posterior probabilities associated with y_a, y_b and y_{ab} respectively. Namely, $\alpha_x = W(x|y_a)$, $\beta_x = W(x|y_b)$, and a short calculation shows that

$$\gamma_x = Q(x|y_{ab}) = \frac{\pi_a \alpha_x + \pi_b \beta_x}{\pi_{ab}} = \frac{\pi_a \alpha_x + \pi_b \beta_x}{\pi_a + \pi_b} . \tag{4}$$

Thus, after canceling terms, one gets that

$$\Delta I = I(W) - I(Q) = \sum_{x \in \mathcal{X}} \Delta I_x , \tag{5}$$

where $\Delta I_x \triangleq \pi_{ab}\eta(\gamma_x) - \pi_a\eta(\alpha_x) - \pi_b\eta(\beta_x)$.

In order to bound ΔI, we give two bounds on ΔI_x. The first bound was derived in [5],

$$\Delta I_x \leq (\pi_a + \pi_b) \cdot d_1(\alpha_x, \beta_x) , \tag{6}$$

where for $\alpha \geq 0$ and $\zeta \in \mathbb{R}$, we define $d_1(\alpha, \zeta) \triangleq |\zeta - \alpha|$.

The subscript "1" in d_1 is suggestive of the L_1 distance. Note that we will generally use α to denote a probability associated with an input letter, while ζ will denote a "free" real variable, possibly negative. We will keep to this convention for the vector case as well. In addition, let us point out that the bound in (6) was derived assuming a uniform input distribution, however remains valid for the general case.

We now derive the second bound on ΔI_x. For the case where $\alpha_x, \beta_x > 0$,

$$\begin{aligned}
\Delta I_x &= \pi_a(\eta(\gamma_x) - \eta(\alpha_x)) + \pi_b(\eta(\gamma_x) - \eta(\beta_x)) \\
&\overset{(a)}{\leq} \pi_a\eta'(\alpha_x)(\gamma_x - \alpha_x) + \pi_b\eta'(\beta_x)(\gamma_x - \beta_x) \\
&\overset{(b)}{=} \frac{\pi_a\pi_b}{\pi_a + \pi_b}(\alpha_x - \beta_x)(\eta'(\beta_x) - \eta'(\alpha_x)) \\
&\overset{(c)}{\leq} \frac{1}{4}(\pi_a + \pi_b)(\alpha_x - \beta_x)^2(-\eta''(\lambda)) ,
\end{aligned}$$

where in (a) we used the concavity of $\eta(\cdot)$, in (b) the definition of γ_x (see (4)), and in (c) the AM-GM inequality and the mean value theorem where $\lambda = \theta\alpha_x + (1-\theta)\beta_x$ for some $\theta \in [0,1]$. Using the monotonicity of $-\eta''(p) = 1/p$ we get $-\eta''(\lambda) \leq 1/\min(\alpha_x, \beta_x)$. Thus,

$$\Delta I_x \leq (\pi_a + \pi_b) \cdot d_2(\alpha_x, \beta_x) , \tag{7}$$

where

$$d_2(\alpha, \zeta) \triangleq \begin{cases} \frac{(\zeta - \alpha)^2}{\min(\alpha, \zeta)} & \alpha, \zeta > 0 , \\ \infty & \text{otherwise} . \end{cases}$$

The subscript "2" in d_2 is suggestive of the squaring in the numerator. Combining (6) and (7) yields

$$\Delta I_x \leq (\pi_a + \pi_b) \cdot d(\alpha_x, \beta_x) \ , \qquad (8)$$

where

$$d(\alpha, \zeta) \triangleq \min(d_1(\alpha, \zeta), d_2(\alpha, \zeta)) \ . \qquad (9)$$

Returning to (5) using (8) we get

$$\Delta I \leq (\pi_a + \pi_b)|\mathcal{X}| \cdot d(\boldsymbol{\alpha}, \boldsymbol{\beta}) \ , \qquad (10)$$

where

$$d(\boldsymbol{\alpha}, \boldsymbol{\zeta}) \triangleq \max_{x \in \mathcal{X}} d(\alpha_x, \zeta_x) \ . \qquad (11)$$

We note that we use $\max$ in (11) instead of a summation to simplify the upcoming derivations. Moreover, according to (10), it suffices to show the existence of a pair that is "close" in the sense of d, assuming that π_a, π_b are also small enough.

Since we are interested in lowering the right hand side of (10), we limit our search to a subset of $\mathcal{Y}$, as was done in [5]. Namely, $\mathcal{Y}_{\text{small}} \triangleq \{y \in \mathcal{Y} : \pi(y) \leq 2/|\mathcal{Y}|\}$, which implies

$$|\mathcal{Y}_{\text{small}}| \geq \frac{|\mathcal{Y}|}{2} \ . \qquad (12)$$

Hence, $\pi_a + \pi_b \leq 4/|\mathcal{Y}|$ and

$$\Delta I \leq \frac{4|\mathcal{X}|}{|\mathcal{Y}|} \cdot d(\boldsymbol{\alpha}, \boldsymbol{\beta}) \ . \qquad (13)$$

We still need to prove the existence of a pair $y_a, y_b \in \mathcal{Y}_{\text{small}}$ that is "close" in the sense of d. To that end, as in [5], we would like to use a sphere-packing approach. A typical use of such an argument assumes a proper metric, yet d is not a metric. Specifically, the triangle-inequality does not hold. The absence of a triangle-inequality is a complication that we will overcome, but some care and effort are called for. Broadly speaking, as usually done in sphere-packing, we aim to show the existence of a critical "sphere" radius, $r_{\text{critical}} = r_{\text{critical}}(|\mathcal{X}|, |\mathcal{Y}|) > 0$. Such a critical radius will ensure the existence of $y_a, y_b \in \mathcal{Y}_{\text{small}}$ with corresponding $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ for which $d(\boldsymbol{\alpha}, \boldsymbol{\beta}) \leq r_{\text{critical}}$.

C. Non-intersecting "spheres"

We start by giving explicit equations for the "spheres" corresponding to d_1 and d_2.

Lemma 3. *For $\alpha \geq 0$ and $r > 0$, define the sets $\mathcal{B}_1, \mathcal{B}_2$ as*

$$\mathcal{B}_i(\alpha, r) \triangleq \{\zeta \in \mathbb{R} : d_i(\alpha, \zeta) \leq r\} \ , \quad i \in \{1, 2\} \ .$$

Then,

$$\mathcal{B}_1(\alpha, r) = \{\zeta \in \mathbb{R} : -r \leq \zeta - \alpha \leq r\}$$

and

$$\mathcal{B}_2(\alpha, r)$$
$$= \left\{\zeta \in \mathbb{R} : -\sqrt{r^2/4 + \alpha \cdot r} + r/2 \leq \zeta - \alpha \leq \sqrt{\alpha \cdot r}\right\} \ .$$

Proof. Assume $\zeta \in \mathcal{B}_1(\alpha, r)$. Then ζ satisfies $|\zeta - \alpha| \leq r$, which is equivalent to $-r \leq \zeta - \alpha \leq r$, and we get the desired result for $\mathcal{B}_1(\alpha, r)$. Assume now $\zeta \in \mathcal{B}_2(\alpha, r)$. If $\zeta \geq \alpha$, then

$\min(\alpha, \zeta) = \alpha$, and thus $(\zeta - \alpha)^2/\alpha \leq r$, which implies $0 \leq \zeta - \alpha \leq \sqrt{\alpha \cdot r}$. If $\zeta \leq \alpha$, then $\min(\alpha, \zeta) = \zeta$, and thus, $\frac{(\zeta - \alpha)^2}{\zeta} \leq r$, which implies $-\sqrt{\frac{r^2}{4} + \alpha \cdot r} + \frac{r}{2} \leq \zeta - \alpha \leq 0$. The union of the two yields the desired result for $\mathcal{B}_2(\alpha, r)$. $\blacksquare$

Thus, we define $\mathcal{B}(\alpha, r) \triangleq \{\zeta \in \mathbb{R} : d(\alpha, \zeta) \leq r\}$, and note that $\mathcal{B}(\alpha, r) = \mathcal{B}_1(\alpha, r) \cup \mathcal{B}_2(\alpha, r)$, since d takes the $\min$ of the two distances. Namely,

$$\mathcal{B}(\alpha, r) = \{\zeta \in \mathbb{R} : -\underline{\omega}(\alpha, r) \leq \zeta - \alpha \leq \overline{\omega}(\alpha, r)\} \ , \quad (14)$$

where $\underline{\omega}(\alpha, r) \triangleq \max\left(\sqrt{r^2/4 + \alpha \cdot r} - r/2, r\right)$ and $\overline{\omega}(\alpha, r) \triangleq \max\left(\sqrt{\alpha \cdot r}, r\right)$. To extend $\mathcal{B}$ to vectors, we define $\mathbb{R}^{|\mathcal{X}|}$ as the set of vectors with real entries that are indexed by $\mathcal{X}$,

$$\mathbb{R}^{|\mathcal{X}|} \triangleq \{\boldsymbol{\zeta} = (\zeta_x)_{x \in \mathcal{X}} : \zeta_x \in \mathbb{R}\} \ .$$

The set $\mathbb{K}^{|\mathcal{X}|}$ is defined as the set of vectors from $\mathbb{R}^{|\mathcal{X}|}$ with entries summing to 1,

$$\mathbb{K}^{|\mathcal{X}|} \triangleq \left\{\boldsymbol{\zeta} \in \mathbb{R}^{|\mathcal{X}|} : \sum_{x \in \mathcal{X}} \zeta_x = 1\right\} \ .$$

The set $\mathbb{K}_+^{|\mathcal{X}|}$ is the set of probability vectors. Namely, the set of vectors from $\mathbb{K}^{|\mathcal{X}|}$ with non-negative entries,

$$\mathbb{K}_+^{|\mathcal{X}|} \triangleq \left\{\boldsymbol{\zeta} \in \mathbb{K}^{|\mathcal{X}|} : \zeta_x \geq 0\right\} \ .$$

We can now define $\mathcal{B}(\boldsymbol{\alpha}, r)$. For $\boldsymbol{\alpha} \in \mathbb{K}_+^{|\mathcal{X}|}$ let

$$\mathcal{B}(\boldsymbol{\alpha}, r) \triangleq \left\{\boldsymbol{\zeta} \in \mathbb{R}^{|\mathcal{X}|} : d(\boldsymbol{\alpha}, \boldsymbol{\zeta}) \leq r\right\} \ . \qquad (15)$$

Using (11) and (14) we have a simple characterization of $\mathcal{B}(\boldsymbol{\alpha}, r)$ as a box: a Cartesian product of segments. That is,

$$\mathcal{B}(\boldsymbol{\alpha}, r) = \Big\{\boldsymbol{\zeta} \in \mathbb{R}^{|\mathcal{X}|} :$$
$$-\underline{\omega}(\alpha_x, r) \leq \zeta_x - \alpha_x \leq \overline{\omega}(\alpha_x, r)\Big\} \ . \qquad (16)$$

We stress that the box $\mathcal{B}(\boldsymbol{\alpha}, r)$ contains $\boldsymbol{\alpha}$, but is not necessarily centered at it.

Recall our aim is finding an r_{critical}. Using our current notation, r_{critical} must imply the existence of distinct $y_a, y_b \in \mathcal{Y}_{\text{small}}$ such that $\boldsymbol{\beta} \in \mathcal{B}(\boldsymbol{\alpha}, r_{\text{critical}})$. Note that the set $\mathcal{B}(\boldsymbol{\alpha}, r)$ is contained in $\mathbb{R}^{|\mathcal{X}|}$. However, since the boxes are induced by points $\boldsymbol{\alpha}$ in the subspace $\mathbb{K}_+^{|\mathcal{X}|}$ of $\mathbb{R}^{|\mathcal{X}|}$, the sphere-packing would yield a tighter result if performed in $\mathbb{K}^{|\mathcal{X}|}$ rather than in $\mathbb{R}^{|\mathcal{X}|}$. Then, for $\boldsymbol{\alpha} \in \mathbb{K}_+^{|\mathcal{X}|}$ and $r > 0$, let us define

$$\mathcal{B}_{\mathbb{K}}(\boldsymbol{\alpha}, r) = \mathcal{B}(\boldsymbol{\alpha}, r) \cap \mathbb{K}^{|\mathcal{X}|} \ . \qquad (17)$$

When considering $\mathcal{B}_{\mathbb{K}}(\boldsymbol{\alpha}, r)$ in place of $\mathcal{B}(\boldsymbol{\alpha}, r)$, we have gained in that the affine dimension (see [9, Section 2.1.3]) of $\mathcal{B}_{\mathbb{K}}(\boldsymbol{\alpha}, r)$ is $|\mathcal{X}| - 1$ while that of $\mathcal{B}(\boldsymbol{\alpha}, r)$ is $|\mathcal{X}|$. However, we have lost in simplicity: the set $\mathcal{B}_{\mathbb{K}}(\boldsymbol{\alpha}, r)$ is not a box. Indeed, a moment's thought reveals that any subset of $\mathbb{K}^{|\mathcal{X}|}$ with more than one element cannot be a box.

We now show how to overcome the above loss. That is, we show a subset of $\mathcal{B}_{\mathbb{K}}(\boldsymbol{\alpha}, r)$ which is — up to a simple transform

— a box. Denote the index of the largest entry of a vector $\boldsymbol{\alpha} \in \mathbb{K}^{|\mathcal{X}|}$ as $x_{\max}(\boldsymbol{\alpha})$, namely, $x_{\max}(\boldsymbol{\alpha}) \triangleq \arg\max_{x \in \mathcal{X}} \alpha_x$.

In case of ties, define $x_{\max}(\boldsymbol{\alpha})$ in an arbitrary yet consistent manner. For $x_{\max} = x_{\max}(\boldsymbol{\alpha})$ given, or clear from the context, define $\boldsymbol{\zeta}'$ as $\boldsymbol{\zeta}$, with index $x_{\max}$ deleted. That is, for a given $\boldsymbol{\zeta} \in \mathbb{K}^{|\mathcal{X}|}$, $\boldsymbol{\zeta}' = (\zeta_x)_{x \in \mathcal{X}'} \in \mathbb{R}^{|\mathcal{X}|-1}$, where $\mathcal{X}' \triangleq \mathcal{X} \setminus \{x_{\max}\}$. Note that for $\boldsymbol{\zeta} \in \mathbb{K}^{|\mathcal{X}|}$, all the entries sum to one. Thus, given $\boldsymbol{\zeta}'$ and $x_{\max}$, we know $\boldsymbol{\zeta}$. Next, for $\boldsymbol{\alpha} \in \mathbb{K}_+^{|\mathcal{X}|}$ and $r > 0$, define the set

$$
\mathcal{C}(\boldsymbol{\alpha}, r) = \{\boldsymbol{\zeta} \in \mathbb{K}^{|\mathcal{X}|} :
$$
$$
\forall x \in \mathcal{X}' , \ -\omega'(\alpha_x, r) \le \zeta_x - \alpha_x \le \omega'(\alpha_x, r)\} , \quad (18)
$$

where $x_{\max} = x_{\max}(\boldsymbol{\alpha})$ and

$$
\omega'(\alpha, r) \triangleq \frac{\omega(\alpha, r)}{|\mathcal{X}| - 1} . \quad (19)
$$

Lemma 4. *Let $\boldsymbol{\alpha} \in \mathbb{K}_+^{|\mathcal{X}|}$ and $r > 0$ be given. Let $x_{\max} = x_{\max}(\boldsymbol{\alpha})$. Then, $\mathcal{C}(\boldsymbol{\alpha}, r) \subset \mathcal{B}_{\mathbb{K}}(\boldsymbol{\alpha}, r)$.*

Proof. It can be easily shown that $0 \le \underline{\omega}(\alpha, r) \le \overline{\omega}(\alpha, r)$. Thus, since (18) holds, it suffices to show that

$$
-\underline{\omega}(\alpha_{x_{\max}}, r) \le \zeta_{x_{\max}} - \alpha_{x_{\max}} \le \underline{\omega}(\alpha_{x_{\max}}, r) . \quad (20)
$$

Indeed, summing the condition in (18) over all $x \in \mathcal{X}'$ gives

$$
\sum_{x \in \mathcal{X}'} -\omega'(\alpha_x, r) \le \sum_{x \in \mathcal{X}'} \zeta_x - \sum_{x \in \mathcal{X}'} \alpha_x \le \sum_{x \in \mathcal{X}'} \omega'(\alpha_x, r) .
$$

Since $\underline{\omega}(\alpha, r)$ is a monotonically non-decreasing function of α, we can simplify the above to

$$
-\underline{\omega}(\alpha_{x_{\max}}, r) \le \sum_{x \in \mathcal{X}'} \zeta_x - \sum_{x \in \mathcal{X}'} \alpha_x \le \underline{\omega}(\alpha_{x_{\max}}, r) .
$$

Since both $\boldsymbol{\zeta}$ and $\boldsymbol{\alpha}$ are in $\mathbb{K}^{|\mathcal{X}|}$, the middle term in the above is $\alpha_{x_{\max}} - \zeta_{x_{\max}}$. Thus, (20) follows. ∎

Recall that our plan is to ensure the existence of a "close" pair by using a sphere-packing approach. However, since the triangle inequality does not hold for d, we must use a somewhat different approach. Towards that end, define the positive quadrant associated with $\boldsymbol{\alpha}$ and r as

$$
\mathcal{Q}'(\boldsymbol{\alpha}, r) = \{\boldsymbol{\zeta}' \in \mathbb{R}^{|\mathcal{X}|-1} :
$$
$$
\forall x \in \mathcal{X}', \ 0 \le \zeta_x - \alpha_x \le \omega'(\alpha_x, r)\} ,
$$

where $x_{\max} = x_{\max}(\boldsymbol{\alpha})$ and $\omega'(\alpha, r)$ is as defined in (19).

Lemma 5. *Let $y_a, y_b \in \mathcal{Y}$ be such that $x_{\max}(\boldsymbol{\alpha}) = x_{\max}(\boldsymbol{\beta})$. If $\mathcal{Q}'(\boldsymbol{\alpha}, r)$ and $\mathcal{Q}'(\boldsymbol{\beta}, r)$ have a non-empty intersection, then $d(\boldsymbol{\alpha}, \boldsymbol{\beta}) \le r$.*

Proof. By (15), (17), and Lemma 4, it suffices to prove that $\boldsymbol{\beta} \in \mathcal{C}(\boldsymbol{\alpha}, r)$. Define $\mathcal{C}'(\boldsymbol{\alpha}, r)$ as the result of applying a prime operation on each member of $\mathcal{C}(\boldsymbol{\alpha}, r)$, where $x_{\max} = x_{\max}(\boldsymbol{\alpha})$. Hence, we must equivalently prove that $\boldsymbol{\beta}' \in \mathcal{C}'(\boldsymbol{\alpha}, r)$. By (18), we must show that for all $x \in \mathcal{X}'$,

$$
-\omega'(\alpha_x, r) \le \beta_x - \alpha_x \le \omega'(\alpha_x, r) . \quad (21)
$$

Since we know that the intersection of $\mathcal{Q}'(\boldsymbol{\alpha}, r)$ and $\mathcal{Q}'(\boldsymbol{\beta}, r)$ is non-empty, let $\boldsymbol{\zeta}'$ be a member of both sets. Thus, we know that for $x \in \mathcal{X}'$, $0 \le \zeta_x - \alpha_x \le \omega'(\alpha_x, r)$, and $0 \le \zeta_x - \beta_x \le \omega'(\beta_x, r)$. For each $x \in \mathcal{X}'$ we must consider two cases: $\alpha_x \le \beta_x$ and $\alpha_x > \beta_x$.

Consider first the case $\alpha_x \le \beta_x$. Since $\zeta_x - \alpha_x \le \omega'(\alpha_x, r)$ and $\beta_x - \zeta_x \le 0$, we conclude that $\beta_x - \alpha_x \le \omega'(\alpha_x, r)$. Conversely, since $\beta_x - \alpha_x \ge 0$ and, by (19), $\omega'(\alpha_x, r) \ge 0$, we have that $\beta_x - \alpha_x \ge -\omega'(\alpha_x, r)$. Thus we have shown that both inequalities in (21) hold.

To finish the proof, consider the case $\alpha_x > \beta_x$. We have already established that $\omega'(\alpha_x, r) \ge 0$. Thus, since by assumption $\beta_x - \alpha_x \le 0$, we have that $\beta_x - \alpha_x \le \omega'(\alpha_x, r)$. Conversely, since $\zeta_x - \beta_x \le \omega'(\beta_x, r)$ and $\alpha_x - \zeta_x \le 0$, we have that $\alpha_x - \beta_x \le \omega'(\beta_x, r)$. We now recall that by (19), the fact that $\alpha_x \ge \beta_x$ implies that $\omega'(\beta_x, r) \le \omega'(\alpha_x, r)$. Thus, $\alpha_x - \beta_x \le \omega'(\alpha_x, r)$. Negating gives $\beta_x - \alpha_x \ge -\omega'(\alpha_x, r)$, and we have once again proved the two inequalities in (21). ∎

D. Weighted "sphere"-packing

The volume of our "sphere" $\mathcal{Q}'(\boldsymbol{\alpha}, r)$ unfortunately depends on $\boldsymbol{\alpha}$. We would like then to alleviate this dependency by defining a density over $\mathbb{R}^{|\mathcal{X}|-1}$ and derive a lower bound on the weight of $\mathcal{Q}'(\boldsymbol{\alpha}, r)$. Let $\varphi : \mathbb{R} \to \mathbb{R}$ be defined as $\varphi(\zeta) \triangleq 1/\sqrt{4\zeta}$. Next, for $\boldsymbol{\zeta}' \in \mathbb{R}^{|\mathcal{X}|-1}$, abuse notation and define $\varphi : \mathbb{R}^{|\mathcal{X}|-1} \to \mathbb{R}$ as $\varphi(\boldsymbol{\zeta}') \triangleq \prod_{x \in \mathcal{X}'} \varphi(\zeta_x)$. The weight of $\mathcal{Q}'(\boldsymbol{\alpha}, r)$ is then defined as $M[\mathcal{Q}'(\boldsymbol{\alpha}, r)] \triangleq \int_{\mathcal{Q}'(\boldsymbol{\alpha}, r)} \mathrm{d}\varphi$. The following lemma proposes a lower bound on $M[\mathcal{Q}'(\boldsymbol{\alpha}, r)]$ that does not depend on $\boldsymbol{\alpha}$.

Lemma 6. *The weight $M[\mathcal{Q}'(\boldsymbol{\alpha}, r)]$ satisfies*

$$
M[\mathcal{Q}'(\boldsymbol{\alpha}, r)] \ge r^{\frac{|\mathcal{X}|-1}{2}} \left(\sqrt{2 + \frac{1}{|\mathcal{X}| - 1}} - \sqrt{2} \right)^{|\mathcal{X}|-1} . \quad (22)
$$

Proof. Since $\varphi(\boldsymbol{\zeta}')$ is a product,

$$
M[\mathcal{Q}'(\boldsymbol{\alpha}, r)] = \prod_{x \in \mathcal{X}'} \int_{\alpha_x}^{\alpha_x + \omega'(\alpha_x, r)} \frac{\mathrm{d}\zeta_x}{2\sqrt{\zeta_x}}
$$
$$
= \prod_{x \in \mathcal{X}'} \psi_r(\alpha_x) ,
$$

where $\psi_r(\alpha) \triangleq \sqrt{\alpha + \omega'(\alpha, r)} - \sqrt{\alpha}$. It can be shown that $\psi_r(\alpha)$ is decreasing when $\alpha < 2r$ simply by using the first derivative. As for $\alpha \ge 2r$, it can be shown that $\psi_r'(\alpha)$ is non-zero. Since $\psi_r'(2r) > 0$ we conclude that $\psi_r(\alpha)$ is increasing. By continuity we conclude that $\psi_r(\alpha)$ is minimal for $\alpha = 2r$ and thus we get (22). ∎

We divide the letters in $\mathcal{Y}_{\text{small}}$ to $|\mathcal{X}|$ subsets, according to their $x_{\max}$ value. The largest subset is denoted by $\mathcal{Y}'$, and we henceforth fix $x_{\max}$ accordingly. We limit our search to $\mathcal{Y}'$.

Let $\mathcal{V}'$ be the union of all the quadrants corresponding to possible choices of $\boldsymbol{\alpha}$. Namely,

$$
\mathcal{V}' \triangleq \bigcup_{\substack{\boldsymbol{\alpha} \in \mathbb{K}_+^{|\mathcal{X}|} , \\ x_{\max}(\boldsymbol{\alpha}) = x_{\max}}} \mathcal{Q}'(\boldsymbol{\alpha}, r) .
$$

In order to bound the weight of $\mathcal{V}'$, we introduce the simpler set $\mathcal{U}'$.

$$\mathcal{U}' \triangleq \left\{ \boldsymbol{\zeta}' \in \mathbb{R}^{|\mathcal{X}|-1} : \sum_{x \in \mathcal{X}'} \zeta_x \leq 2, \ \zeta_x \geq 0 \ \forall x \in \mathcal{X}' \right\} .$$

The constraint $r \leq 1$ in the following lemma will be motivated shortly.

Lemma 7. *Let $r \leq 1$. Then, $\mathcal{V}' \subseteq \mathcal{U}'$.*

Proof. Assume $\boldsymbol{\zeta}' \in \mathcal{V}'$. Then, there exists $\boldsymbol{\alpha} \in \mathbb{K}_+^{|\mathcal{X}|}$ such that $0 \leq \zeta_x - \alpha_x \leq \omega'(\alpha_x, r)$ for all $x \in \mathcal{X}'$. Hence, $\zeta_x \geq 0$ for all $x \in \mathcal{X}'$. Moreover,

$$\sum_{x \in \mathcal{X}'} \zeta_x \leq \sum_{x \in \mathcal{X}'} \alpha_x + \sum_{x \in \mathcal{X}'} \omega'(\alpha_x, r)$$
$$\leq 1 - \alpha_{x_{\max}} + \underline{\omega}(\alpha_{x_{\max}}, r) . \quad (23)$$

There are two case to consider. In the case where $\alpha_{x_{\max}} \geq 2r$ we have

$$\sum_{x \in \mathcal{X}'} \zeta_x \leq 1 - \alpha_{x_{\max}} + \sqrt{\frac{r^2}{4} + \alpha_{x_{\max}} r} - \frac{r}{2}$$
$$\leq 1 - \alpha_{x_{\max}} + \sqrt{\frac{\alpha_{x_{\max}}^2}{16} + \frac{\alpha_{x_{\max}}^2}{2}} - \frac{r}{2}$$
$$\leq 2 ,$$

where the second inequality is due to the assumption $\alpha_{x_{\max}} \geq 2r$. In the case where $\alpha_{x_{\max}} \leq 2r$, (23) becomes

$$\sum_{x \in \mathcal{X}'} \zeta_x \leq 1 - \alpha_{x_{\max}} + r \leq 2 - \alpha_{x_{\max}} \leq 2 ,$$

where we assumed $r \leq 1$. Therefore, $\boldsymbol{\zeta}' \in \mathcal{U}'$. ∎

The lemma above and the non-negativity of φ, enable us to upper bound the weight of $\mathcal{V}'$, denoted by $M[\mathcal{V}']$, using $M[\mathcal{V}'] \triangleq \int_{\mathcal{V}'} \mathrm{d}\varphi \leq \int_{\mathcal{U}'} \mathrm{d}\varphi$. We define the mapping $\rho_x = \sqrt{\zeta_x}$ for all $x \in \mathcal{X}'$ and perform a change of variables. As a result, $\mathcal{U}'$ is mapped to

$$\mathcal{S}' \triangleq \left\{ \boldsymbol{\rho}' \in \mathbb{R}^{|\mathcal{X}|-1} : \sum_{x \in \mathcal{X}'} \rho_x^2 \leq 2, \ \rho_x \geq 0 \right\} ,$$

which is a quadrant of a $|\mathcal{X}| - 1$ dimensional ball of a $\sqrt{2}$ radius. The density function φ transforms into the unit uniform density function since $\mathrm{d}\zeta_x / \sqrt{4\zeta_x} = \mathrm{d}\rho_x$. Hence, for $r \leq 1$,

$$M[\mathcal{V}'] \leq \int_{\mathcal{S}'} \mathrm{d}V = \left(\frac{\pi}{2}\right)^{\frac{|\mathcal{X}|-1}{2}} \frac{1}{\Gamma\left(1 + \frac{|\mathcal{X}|-1}{2}\right)} , \quad (24)$$

where we have used the well known expression for the volume of a multidimensional ball. Finally, we prove Theorem 1.

Proof of Theorem 1. Recall that we are assuming $|\mathcal{Y}| > 2|\mathcal{X}|$. According to the definition of $\mathcal{Y}'$, we get by (12) that

$$|\mathcal{Y}'| \geq \frac{|\mathcal{Y}_{\mathrm{small}}|}{|\mathcal{X}|} \geq \frac{|\mathcal{Y}|}{2|\mathcal{X}|} > 1 . \quad (25)$$

As a result, we have at least two points in $\mathcal{Y}'$, and are therefore in a position to apply a sphere-packing argument. Towards this end, let r be such that the starred equality in the following derivation holds:

$$\sum_{\boldsymbol{\alpha} \in \mathcal{Y}'} M[\mathcal{Q}'(\boldsymbol{\alpha}, r)]$$
$$\geq \frac{|\mathcal{Y}|}{2|\mathcal{X}|} \cdot r^{\frac{|\mathcal{X}|-1}{2}} \left(\sqrt{2 + \frac{1}{|\mathcal{X}|-1}} - \sqrt{2}\right)^{|\mathcal{X}|-1}$$
$$\stackrel{(*)}{=} \left(\frac{\pi}{2}\right)^{\frac{|\mathcal{X}|-1}{2}} \frac{1}{\Gamma\left(1 + \frac{|\mathcal{X}|-1}{2}\right)} \quad (26)$$
$$\geq M[\mathcal{V}'] .$$

Namely,

$$r \triangleq \frac{\pi}{4} \left(\sqrt{1 + \frac{1}{2(|\mathcal{X}|-1)}} - 1\right)^{-2}$$
$$\cdot \left(\frac{2|\mathcal{X}|}{\Gamma\left(1 + \frac{|\mathcal{X}|-1}{2}\right)}\right)^{\frac{2}{|\mathcal{X}|-1}} \cdot |\mathcal{Y}|^{-\frac{2}{|\mathcal{X}|-1}} . \quad (27)$$

There are two cases to consider. If $r \leq 1$, then all of (26) holds, by (22), (24) and (25). We take $r_{\mathrm{critical}} = r$, and deduce the existence of a pair $y_a, y_b \in \mathcal{Y}'$ for which $d(\boldsymbol{\alpha}, \boldsymbol{\beta}) \leq r$. Indeed, assuming otherwise would contradict (26), since each $\mathcal{Q}'$ in the sum is contained in $\mathcal{V}'$, and, by Lemma 5 and our assumption, all summed $\mathcal{Q}'$ are disjoint.

We next consider the case $r > 1$. Now, any pair of letters $y_a, y_b \in \mathcal{Y}'$ satisfies $d(\boldsymbol{\alpha}, \boldsymbol{\beta}) \leq r$. Indeed, by (9) and (11),

$$d(\boldsymbol{\alpha}, \boldsymbol{\beta}) \leq \|\boldsymbol{\alpha} - \boldsymbol{\beta}\|_\infty \leq 1 < r ,$$

where $\|\cdot\|_\infty$ is the maximum norm.

We have proved the existence of $y_a, y_b \in \mathcal{Y}' \subset \mathcal{Y}_{\mathrm{small}}$ for which $d(\boldsymbol{\alpha}, \boldsymbol{\beta}) \leq r$. By (13) and (27), the proof is finished. ∎

REFERENCES

[1] E. Arıkan, "Channel polarization: A method for constructing capacity-achieving codes for symmetric binary-input memoryless channels," *IEEE Trans. Inf. Theory*, vol. 55, no. 7, pp. 3051–3073, July 2009.

[2] I. Tal and A. Vardy, "How to construct polar codes," *IEEE Trans. Inf. Theory*, vol. 59, no. 10, pp. 6562–6582, October 2013.

[3] R. Pedarsani, S. H. Hassani, I. Tal, and E. Telatar, "On the construction of polar codes," in *2011 IEEE Int'l Symp. on Inf. Theory (ISIT)*, July 2011, pp. 11–15.

[4] I. Tal, A. Sharov, and A. Vardy, "Constructing polar codes for non-binary alphabets and macs," in *2012 IEEE Int'l Symp. on Inf. Theory (ISIT)*, July 2012, pp. 2132–2136.

[5] T. C. Gulcu, M. Ye, and A. Barg, "Construction of polar codes for arbitrary discrete memoryless channels," in *2016 IEEE Int'l Symp. on Inf. Theory (ISIT)*, July 2016, pp. 51–55.

[6] U. Pereg and I. Tal, "Channel upgradation for non-binary input alphabets and MACs," *To appear in IEEE Trans. Inform. Theory*, 2017.

[7] I. Tal, "On the construction of polar codes for channels with moderate input alphabet sizes," in *2015 IEEE Int'l Symp. on Inf. Theory (ISIT)*, June 2015, pp. 1297–1301.

[8] B. M. Kurkoski and H. Yagi, "Quantization of binary-input discrete memoryless channels," *IEEE Trans. Inf. Theory*, vol. 60, no. 8, pp. 4544–4552, August 2014.

[9] S. Boyd and L. Vandenberghe, *Convex optimization*. Cambridge university press, 2004.