Information-Distilling Quantizers

Bobak Nazer, Or Ordentlich, and Yury Polyanskiy

***Abstract*—Let X and Y be dependent random variables. We consider the problem of designing a scalar quantizer for Y to maximize the mutual information between its output and X, and study fundamental properties and bounds for this form of quantization. Our main focus is the regime of low $I(X;Y)$, where we show that for a binary X, there always exists an M-level quantizer attaining mutual information of $\Omega(-M \cdot I(X;Y)/\log(I(X;Y))$ and that there exists pairs of X,Y for which the mutual information attained by any M-level quantizer is $\mathcal{O}(-M \cdot I(X;Y)/\log(I(X;Y)))$.**

I. Introduction

Quantization plays a central role in many information processing systems. For instance, when the data comes from a continuous alphabet, quantization is a pre-requisite for digital processing. However, even if the data comes from a discrete alphabet, reducing its cardinality often leads to more efficient processing.

Let X and Y be a pair of random variables with a given distribution P_{XY}. This paper deals with the problem of quantizing Y into $M < |\mathcal{Y}|$ values, under the objective of maximizing the mutual information between the quantizer's output and X. Thus, the optimal quantizer under this setup is

$$\underset{f:\mathcal{Y}\to[M]}{\operatorname{argsup}}\ I(X;f(Y)), \tag{1}$$

where $[M] \triangleq \{1,2,\ldots,M\}$. We will use the following shorthand[1] to denote the value of the mutual information attained by the optimal M-ary quantizer.

$$I(X;[Y]_M) \triangleq \sup_{\tilde{Y}\in[Y]_M} I(X;\tilde{Y}). \tag{2}$$

where $[Y]_M$ is the set of all (deterministic) M-ary quantizations of Y,

$$[Y]_M \triangleq \{f(Y)\ :\ f:\mathcal{Y}\to[M]\}.$$

When X and Y are thought of as the input and output of a channel, respectively, the problem (1) boils down to designing the M-level quantizer that maximizes the information rate, whereas (2) is the highest information rate attainable. It is therefore not surprising that this problem has received considerable attention.[2] For example, it is well known [2, Section 2.11] that when X is a BPSK input to an AWGN channel with output Y it holds that $I(X;[Y]_2) \geq 2I(X;Y)/\pi$ and this is achieved by taking $f(\cdot)$ to be the maximum a posteriori (MAP) estimator of X from Y.[3]

A characterization of (2) is also required for practically constructing polar codes, since the large output cardinality of polarized channels makes it challenging to evaluate their respective capacities (and identify "frozen" bits). Efficient techniques for channel output quantization that preserve mutual information were developed to overcome this obstacle, and played a major role in the process of making polar codes implementable [4]–[6]. Specifically, it was recently shown in [6] that, for arbitrary P_{XY}, it holds that $I(X;Y) - I(X;[Y]_M) \leq \mathcal{O}(M^{-2/(|\mathcal{X}|-1)})$. The works [4]–[6], among others, also provided polynomial complexity sub-optimal algorithms for designing such quantizers. In addition, for binary X, an algorithm for determining the optimal quantizer was proposed in [7] that runs in time $\mathcal{O}(|\mathcal{Y}|^3)$.

In this paper, we ignore the algorithmic aspects of finding the optimal M-level quantizer and instead focus on the fundamental properties of the function $I(X;[Y]_M)$. In particular, our main interest is in identifying the joint distributions P_{XY} that are the most difficult to quantize, and in the value of $I(X;[Y]_M)$ for these cases. Special attention will be given to the binary case, where $X \sim \mathrm{Bernoulli}(p)$ for some p. In this setting, it may seem at a first glance that the optimal binary quantizer should always retain a significant fraction of $I(X;Y)$, and that the MAP quantizer should be sufficient to this end. For large $I(X;Y)$, this is indeed the case, as we show in Proposition 5. This is also the case for the binary AWGN channel for all values of $I(X;Y)$, since the MAP quantizer always retains at least $2/\pi \approx 63.66\%$ of the mutual information.

We state our main result next, with proof deferred to Section III-C.[4]

Theorem 1: If $X \sim \mathrm{Bernoulli}(1/2)$ and $I(X;Y) = \beta > 0$,

BN is with the Department of Electrical and Computer Engineering, Boston University. e-mail: bobak@bu.edu. OO is with the Department of Electrical and Computer Engineering, Boston University and the Department of Electrical Engineering and Computer Science, MIT. e-mail: ordent@mit.edu. YP is with the Department of Electrical Engineering and Computer Science, MIT. e-mail: yp@mit.edu.

The work was supported in part by the NSF CAREER award CCF-12-53205, the Center for Science of Information (CSoI), an NSF Science and Technology Center, under grant agreement CCF-09-39370, and NSF grant CCF-1618800.

[1] This notation is meant to suggest the distance from a point to a set.

[2] This problem is also connected to the log-loss distortion criterion and the information bottleneck tradeoff [1]. An in-depth discussion will appear in an extended version.

[3] It was recently demonstrated in [3] that, if instead of BPSK, an asymmetric signaling scheme is used, the AWGN capacity can be attained at low SNR with an asymmetric 2-level quantizer.

[4] Logarithms are generally taken w.r.t. base 2 in this paper, with the exception of the $\ln$ function that is taken w.r.t. base e.

we have for binary quantization

$$I(X;[Y]_2) \geq \frac{1}{3e} \frac{\beta}{1+\ln\left(\frac{1}{\beta}\right)}. \tag{3}$$

Furthermore, for any $\eta \in (0,1)$ and any natural $M < 12\max\{\log\left(\frac{1}{\beta}\right),1\}/(1-\eta)^2$

$$I(X;[Y]_M) \geq (M-1)\frac{\beta}{\max\{\log\left(\frac{1}{\beta}\right),1\}}\frac{\eta(1-\eta)^2}{12}. \tag{4}$$

Finally, for any $0<\beta\leq 1$, there exist distributions P_{XY} with $X \sim \text{Bernoulli}(1/2)$ and $I(X;Y)=\beta$, for which

$$I(X;[Y]_M) \leq 2M\frac{\beta}{\ln\left(\frac{e\log(e)}{2\beta}\right)}, \tag{5}$$

for every natural M.

Note that this is in stark contrast to the intuition from the binary AWGN channel. While for the former, two quantization levels suffice for retaining a $2/\pi$ fraction of $I(X;Y)$, Theorem 1 shows that there exist distributions for which at least $\Omega(\log(1/I(X;Y)))$ quantization levels are needed in order to retain a fixed fraction of $I(X;Y)$. Furthermore, as illustrated in Section III, for small $I(X;Y)$ and $M=2$, the MAP quantizer can be arbitrary bad w.r.t. the optimal quantizer, which is in general not "symmetric".

For a fixed distribution P_X on $\mathcal{X}$, we define and study the "information distillation" function

$$\text{ID}_M(P_X,\beta) \triangleq \inf_{P_{Y|X}\,:\,I(X;Y)\geq\beta} I(X;[Y]_M), \tag{6}$$

where the infimum is taken w.r.t. to all channels with input alphabet $\mathcal{X}$ and arbitrary (possibly continuous) output alphabet such that the mutual information is at least β. With this notation, Theorem 1 states that $\text{ID}_M(\text{Bernoulli}(1/2),\beta) = \Theta\left(M\beta/\log(1/\beta)\right)$, and in fact, as briefly argued in Section III-C, the same scaling law continues to hold for $\text{ID}_M(\text{Bernoulli}(p),\beta)$, $0<p<1$.

As discussed above, prior work [4]–[6] has focused on bounding the additive gap. In our notation, this corresponds to bounding $\Delta I_M^* \triangleq \sup_{\beta,P_X}\beta - \text{ID}_M(P_X,\beta)$. In particular, the bound derived in [6] on ΔI_M^* is equivalent to the following "constant-gap" result: for every P_X, $\text{ID}_M(P_X,\beta) \geq \beta - \nu(|\mathcal{X}|)M^{-2/(|\mathcal{X}|-1)}$ for some function ν. For small β, however, results of this form are less informative. Indeed, for binary-input channels and small β, this bound requires M to scale like $\beta^{-1/2}$ in order to preserve a constant fraction of the mutual information. On the other hand, our result shows that $M=\mathcal{O}(\log(1/\beta))$ suffices for binary-input channels.

II. Properties of $I(X;[Y]_M)$ and $\text{ID}_M(P_X,\beta)$

Let P_{XY} be a joint distribution on $\mathcal{X}\times\mathcal{Y}$ and consider the function $I(X;[Y]_M)$, as defined in (2). The restriction to deterministic functions incurs no loss of generality, see e.g., [7]. Indeed, any random function of y, can be expressed as $f(y,U)$ where U is some random variable statistically independent of (X,Y). Thus,

$$I(X;f(Y,U)) \leq I(X;f(Y,U),U) = I(X;f(Y,U)|U)$$

and hence there must exist some u for which $I(X;f(Y,u)) \geq I(X;f(Y;U))$. Furthermore, for any function $f:\mathcal{Y}\to[M]$, we can associate a disjoint partition of the cube $[0,1]^{|\mathcal{X}|}$ into M regions $\mathcal{I}_1,\ldots,\mathcal{I}_M$, such that $f(y)=i$ iff $P_{X|Y=y}\in\mathcal{I}_i$ for $i=1,\ldots,M$. A remarkable result of Burshtein et al. [8, Theorem 1] shows that the supremum in (2) can w.l.o.g. be restricted to functions for which there exists an associated partition where the regions $\mathcal{I}_1,\ldots,\mathcal{I}_M$ are all convex.

Below, we state simple upper and lower bounds on $I(X;[Y]_M)$. (The proofs of all propositions are left to an extended version due to space constraints.)

Proposition 1: For any distribution P_{XY} on $\mathcal{X}\times\mathcal{Y}$ with a finite output alphabet, and $M<|\mathcal{Y}|$,

$$\frac{M-1}{|\mathcal{Y}|}I(X;Y) \leq I(X;[Y]_M) \leq \min\{I(X;Y),\log(M)\}.$$

For $K<M$, we can construct a (possibly sub-optimal) K-level quantizer by first finding the optimal M-level quantizer and then quantizing its output to K-levels. This together with the lower bound in Proposition 1, yields the following.

Corollary 1: For natural numbers $K<M$ we have

$$I(X;[Y]_K) \geq \frac{K-1}{M}I(X;[Y]_M).$$

When $X-Y-V$ form a Markov chain in this order, we can simulate any function of V from Y. Thus,

Proposition 2 (Data processing inequality): If $X-Y-V$ form a Markov chain in this order, then

$$I(X;[V]_M) \leq I(X;[Y]_M).$$

Since the supremum of convex functions is convex, we have

Proposition 3: For a fixed P_X, the function $P_{Y|X} \mapsto I(X;[Y]_M)$ is convex.

Remark 1: It is tempting to expect that $I(X;[Y]_M)$ will have "diminishing returns" in M for any P_{XY}, i.e., that it will satisfy the inequality $I(X;[Y]_{M_1\cdot M_2}) \leq I(X;[Y]_{M_1}) + I(X;[Y]_{M_2})$. However, as demonstrated by the following example, this is not the case. Let $X\sim\text{Uniform}(\{0,1,2,3\})$ and $Y=[X+Z] \bmod 4$, where Z is additive noise statistically independent of X with $\Pr(Z=0)=\delta$ and $\Pr(Z=1)=\Pr(Z=2)=\Pr(Z=3)=(1-\delta)/3$. Clearly,

$$I(X;[Y]_4) = I(X;Y) = 2-h(\delta)-(1-\delta)\log(3),$$

and it can be verified that

$$I(X;[Y]_2) = \begin{cases} h\left(\frac{1}{4}\right)-\frac{1}{4}h(\delta)-\frac{3}{4}h\left(\frac{1-\delta}{3}\right) & \delta\leq 1/4, \\ 1-h\left(\frac{1+2\delta}{3}\right) & \delta>1/4. \end{cases}$$

Thus, for this example we have that $2I(X;[Y]_2) < I(X;[Y]_4)$ for all $\delta\notin\{1/4,1\}$.

Remark 2: Without putting any restriction on $|\mathcal{X}|$, the solution of (1) is in general NP-hard, as for $M=2$ and $Y=X$ it reduces to the NP-hard *subset sum problem* [9].

Proposition 4: The function $\mathrm{ID}_M(P_X,\beta)$ is convex and monotonically nondecreasing in β.

III. Bounds for $X\sim\mathrm{Bernoulli}(1/2)$

In this section we provide upper and lower bounds on $\mathrm{ID}_M(P_X,\beta)$, for the special case where $X\sim\mathrm{Bernoulli}(p)$, which we denote by $\mathrm{ID}_M(p,\beta)$. To simplify derivations, we shall further restrict attention to $p=1/2$, though the results we obtain below remain valid for any $0<p<1$ with some correction terms which are qualitatively insignificant. Clearly, for any distribution P_{XY} with $\mathcal{X}=\{1,2\}$ it holds that $I(X;Y)\leq 1$. Thus, β is restricted to the interval $[0,1]$.

A. The Symmetric Quantizer for $M=2$

We begin by analyzing the mutual information induced by the most natural binary quantizer, which is based on the maximum a posteriori (MAP) estimator. In particular, we consider the *symmetric* MAP quantizer

$$f_{\mathrm{MAP}}(y)=\begin{cases}1 & \text{if } \Pr(X=1|Y=y)>1/2\\ 2 & \text{if } \Pr(X=1|Y=y)<1/2\\ \mathrm{Bernoulli}(1/2) & \text{if } \Pr(X=1|Y=y)=1/2\end{cases},$$

where tie-breaking is done via flipping a coin. While the fashion in which ties are broken has no effect on the average error probability, as we shall see, it may effect the mutual information induced by the quantizer. Let $P_{e,\mathrm{MAP}}(y)\triangleq\Pr(f_{\mathrm{MAP}}(Y)\neq X|Y=y)$ and $P_{e,\mathrm{MAP}}\triangleq\mathbb{E}_Y P_{e,\mathrm{MAP}}(Y)$. By concavity of the binary entropy function $h(p)\triangleq -p\log(p)-(1-p)\log(1-p)$ we have that $h(p)\geq 2p$ for any $0\leq p\leq 1/2$, with equality iff $p\in\{0,1/2\}$. Consequently, $H(X|Y)=\mathbb{E}_Y h(P_{e,\mathrm{MAP}}(Y))\geq 2P_{e,\mathrm{MAP}}$. We therefore have that

$$\begin{aligned}I(X;f_{\mathrm{MAP}}(Y))&=1-\mathbb{E}_Y h\left(\Pr(X\neq f_{\mathrm{MAP}}(Y))\right)\\&\geq 1-h(P_{e,\mathrm{MAP}})\geq 1-h\left(\frac{1-I(X;Y)}{2}\right).\end{aligned}\tag{7}$$

B. The High Mutual Information Regime

From (7), we have $\mathrm{ID}_2(1/2,\beta)\geq 1-h\left((1-\beta)/2\right)$. Furthermore, note that the bound (7) is achieved with equality when $P_{Y|X}$ is the binary erasure channel (BEC). For the BEC, however, ties occur frequently and the symmetric MAP quantizer, which involves randomness, turns out not to be a very good choice. Instead of flipping a coin whenever $y=?$, we can always assign a fixed value, say $f(?)=1$, to it. This deterministic asymmetric quantizer, which is also a MAP quantizer, is given by

$$f_Z(y)=\begin{cases}1 & \text{if } y\in\{1,?\},\\ 2 & \text{if } y=0,\end{cases}$$

and induces a Z-channel from X to $f_Z(Y)$ with mutual information

$$I(X;f_Z(Y))=\frac{\beta}{2}h\left(\frac{1-\beta}{2-\beta}\right)+1-h\left(\frac{1-\beta}{2-\beta}\right).\tag{8}$$

It is easy to verify that $f_Z(y)$ is the optimal 1-bit quantizer for the BEC.

We have therefore established the following proposition.

Proposition 5: For all $0\leq\epsilon\leq 1$ we have

$$1-h\left(\frac{\epsilon}{2}\right)\leq\mathrm{ID}_2(1/2,1-\epsilon)\leq 1-\frac{1+\epsilon}{2}h\left(\frac{\epsilon}{1+\epsilon}\right).$$

Thus, for large β, the loss for quantizing the output to one bit is small and the fraction of the mutual information that can be retained approaches 1 as the mutual information increases. In particular, the natural symmetric MAP quantizer is never too bad, and retains a significant fraction of at least $1-h((1-\beta)/2)$ of the mutual information β.

C. The Low Mutual Information Regime

In the small β regime, we arrive at qualitatively different behavior. Consider again the BEC with capacity β for $\beta\ll 1$. By (7) (which becomes an equality for the BEC) and (8), we have that

$$I(X;f_{\mathrm{MAP}}(Y))=1-h\left(\frac{1-\beta}{2}\right)=\frac{\log e}{2}\beta^2+o(\beta^2).\tag{9}$$

$$I(X;f_Z(Y))=\frac{\beta}{2}h\left(\frac{1-\beta}{2-\beta}\right)+1-h\left(\frac{1-\beta}{2-\beta}\right)=\frac{\beta}{2}+o(\beta).\tag{10}$$

Thus, the asymmetric quantizer $f_Z(y)$ retains 50% of the mutual information, whereas the fraction of mutual information retained by the symmetric MAP quantizer vanishes as $\beta\to 0$.

The BEC example may lead to the erroneous conclusion that the weakness of the symmetric MAP quantizer is due to the random tie-breaking, and that a MAP quantizer that breaks ties more cleverly can always retain a significant fraction of $I(X;Y)$. However, this is not the case. To see this consider a channel with binary input and output alphabet $\mathcal{Y}=\{1,2\}\times\{g,b\}$, defined by

$$\Pr(Y=y|X=x)=\begin{cases}\beta & \text{if } y=(x,g)\\ (1-\beta)\left(\frac{1}{2}+\delta\right) & \text{if } y=(x,b)\\ (1-\beta)\left(\frac{1}{2}-\delta\right) & \text{if } y=(1-x,b)\end{cases},$$

for some $0\leq\beta\leq 1$ and $0\leq\delta\leq 1/2$. Note that for $\delta=0$, this channel becomes a BEC with capacity $1-\beta$, but for any $\delta\neq 0$ ties never occur. For any $\delta>0$, the MAP quantizer is therefore unique and deterministic, but as $\delta\to 0$, the channel approaches a BEC, and its performance becomes closer and closer to (9). Similarly, the performance of a binary quantizer that assigns the same value to both "bad" outputs, i.e., $f(y)=2$ if $y=(0,g)$ and $f(y)=1$ otherwise, which is not a MAP quantizer, approaches (10) as $\delta\to 0$, and is therefore much better than the MAP quantizer.

Next, we prove Theorem 1, which requires the following proposition.

Proposition 6: The function $g(t) = -t\ln(t)$ is monotone increasing in $0 < t < 1/e$ and its inverse restricted to this interval satisfies $\frac{1}{e}\cdot\frac{t}{-\ln(t)} < g^{-1}(t) \leq \frac{t}{-\ln(t)}$.

Proof of lower bounds in Theorem 1. Consider the joint distribution P_{XY}, and for any $y \in \mathcal{Y}$ define $\alpha_y \triangleq \Pr(X = 1|Y = y)$, $\bar{\alpha} \triangleq \mathbb{E}(\alpha_Y) = \frac{1}{2}$ and

$$D_y \triangleq D(P_{X|Y=y}\|P_X) = d\left(\alpha_y \,\|\, \bar{\alpha}\right),$$

where $d(p\|q) \triangleq p\log(p/q) + (1-p)\log((1-p)/(1-q))$ is the binary divergence function. We further define the function

$$\bar{F}(\gamma) \triangleq \Pr(D_Y \geq \gamma),$$

and note that it is non-increasing and satisfies

$$I(X;Y) = \mathbb{E}D_Y = \int_0^{\gamma^*} \bar{F}(\gamma)d\gamma, \tag{11}$$

where $\gamma^* = \max_{y\in\mathcal{Y}} D_y \leq 1$. Let $M = 2L+1$ for some natural number L, let $0 = \gamma_0 \leq \gamma_1 \leq \cdots \leq \gamma_L \leq \gamma_{L+1} = \gamma^*+\delta$, for some arbitrary small $\delta > 0$, and define the following M-level quantizer

$$f(y) = \begin{cases} 0 & d(\alpha_y\|\bar{\alpha}) \leq \gamma_1 \\ -\ell & \alpha_y < \bar{\alpha}, \gamma_\ell \leq d(\alpha_y\|\bar{\alpha}) < \gamma_{\ell+1} \\ \ell & \alpha_y > \bar{\alpha}, \gamma_\ell \leq d(\alpha_y\|\bar{\alpha}) < \gamma_{\ell+1} \end{cases}.$$

We have that for $\ell = 1,\ldots,L$

$$d\left(\mathbb{E}[\alpha_Y|f(Y) = -\ell]\|\bar{\alpha}\right) \geq \gamma_\ell,\ d\left(\mathbb{E}[\alpha_Y|f(Y) = \ell]\|\bar{\alpha}\right) \geq \gamma_\ell$$

and by the definition of $\bar{F}(\gamma)$ we also have

$$\Pr\left(\{f(Y) = -\ell\} \cup \{f(Y) = \ell\}\right) = \bar{F}(\gamma_\ell) - \bar{F}(\gamma_{\ell+1}).$$

Thus,

$$I(X;f(Y)) = \sum_{\ell=-L}^{L} \Pr(f(Y) = \ell)D(P_{X|f(Y)=\ell}\|P_X)$$

$$\geq \sum_{\ell=1}^{L}\left(\bar{F}(\gamma_\ell) - \bar{F}(\gamma_{\ell+1})\right)\gamma_\ell = \sum_{\ell=1}^{L}\bar{F}(\gamma_\ell)(\gamma_\ell - \gamma_{\ell-1}) \tag{12}$$

where in the last equality we used $\gamma_0 = 0$ and $\bar{F}(\gamma_{L+1}) = \bar{F}(\gamma^*+\delta) = 0$. Our goal is therefore to choose the numbers $\{\gamma_\ell\}_{\ell=1}^{L}$ such as to maximize (12).

For the special case of $L = 1$, this reduces to $\gamma_1 = \mathrm{argmax}_\gamma\, \gamma\bar{F}(\gamma)$, and with this choice we have $I(X;f(Y)) = \max_\gamma \gamma\bar{F}(\gamma)$. Thus, $\bar{F}(\gamma) \leq \min\{1, I(X;f(Y))/\gamma\}$. Using the identity (11) with $\gamma^* \leq 1$, this yields

$$I(X;Y) \leq \int_0^{I(X;f(Y))} d\gamma + \int_{I(X;f(Y))}^{1} \frac{I(X;f(Y))}{\gamma}d\gamma \tag{13}$$

$$= I(X;f(Y))\left(1 + \ln\frac{1}{I(X;f(Y))}\right)$$

$$= -e\,\frac{I(X;f(Y))}{e}\ln\left(\frac{I(X;f(Y))}{e}\right).$$

Recalling that $L = 1$ corresponds to a quantizer with $M = 2L+1 = 3$ levels and applying Proposition 6, we have therefore obtained

$$I(X;[Y]_3) \geq e\cdot g^{-1}\left(\frac{I(X;Y)}{e}\right) \geq \frac{1}{e}\cdot\frac{I(X;Y)}{1+\ln\left(\frac{1}{I(X;Y)}\right)}.$$

Now, applying Corollary 1, yields (3).

For a general L, the problem of finding $\{\gamma_\ell\}$ such as to maximize (12) is more difficult. We therefore resort to a possibly suboptimal choice according to the rule

$$\gamma_1 = \epsilon I(X;Y),\ \theta = \gamma_1^{-\frac{1}{L}},\ \gamma_\ell = \gamma_1\cdot\theta^{\ell-1}, \tag{14}$$

for $\ell = 2,\ldots,L,L+1$ and some $0 < \epsilon < 1$ to be specified. Note that this choice guarantees that

$$\gamma_{\ell+1} - \gamma_\ell = \theta\left(\gamma_\ell - \gamma_{\ell-1}\right),\ \ell = 1,\ldots,L.$$

This implies that

$$I(X;Y) = \sum_{\ell=0}^{L}\int_{\gamma_\ell}^{\gamma_{\ell+1}}\bar{F}(\gamma)d\gamma \leq \sum_{\ell=0}^{L}(\gamma_{\ell+1}-\gamma_\ell)\bar{F}(\gamma_\ell)$$

$$= \gamma_1 + \theta\sum_{\ell=1}^{L}(\gamma_\ell - \gamma_{\ell-1})\bar{F}(\gamma_\ell)$$

$$\leq \gamma_1 + \theta I(X;f(Y)).$$

Now, setting $\epsilon = 1/(L+1)$ yields

$$I(X;f(Y)) \geq (I(X;Y))^{\frac{L+1}{L}}\frac{L}{(1+L)^{\frac{L+1}{L}}}$$

$$\geq (I(X;Y))^{\frac{L+1}{L}}\cdot\left(1 - \frac{1}{\sqrt{L}}\right), \tag{15}$$

where the last inequality is valid for every $L \geq 1$.

Substituting in $L = \left\lceil 4\max\left\{\log\left(\frac{1}{I(X;Y)}\right), 1\right\}/(1-\eta)^2\right\rceil$, it follows that

$$I(X;f(Y)) \geq 2^{-(1-\eta)^2/4}\left(\frac{1}{2}+\frac{\eta}{2}\right)I(X;Y) \geq \eta I(X;Y).$$

Since $M = 2L+1$ and $L \geq 4$, it follows that we can guarantee $I(X;f(Y)) \geq \eta I(X;Y)$ if $M = \left\lfloor 12\max\left\{\log\left(\frac{1}{I(X;Y)}\right), 1\right\}/(1-\eta)^2\right\rfloor$ and thus $\mathrm{ID}_M(1/2,\beta) \geq \eta\beta$ for this choice of M as well. For smaller values of M, we can apply Corollary 1 to obtain 4. ∎

Remark 3: The proof above only used the assumption that $X \sim \mathrm{Bernoulli}(1/2)$ (rather than $\mathrm{Bernoulli}(p)$ with general p) in order to bound $\gamma^* \leq 1$. The proof can be easily modified to deal with any p, in which case we have $\gamma^* \leq -\log(\min\{p, 1-p\})$. This will require changing the integration limits in (13), and replacing the choice of θ in (14) with $\theta = (\gamma^*/\gamma_1)^{1/L}$.

Proof of upper bound in Theorem 1. It suffices to provide one distribution P_{XY} with $I(X;Y) \geq \beta$ for which no M-level quantizer achieves mutual information exceeding the RHS of (5). To this end, let $X \sim \mathrm{Bernoulli}(1/2)$ and $Y = (X \oplus Z_T, T)$ be the output of a binary-input memoryless output-symmetric (BMS) whose input is X, where T is a mixed random variable in $[0,1/2)$ whose probability density function is given by

$$f_T(t)=\begin{cases} r\delta(t)+\frac{4r}{(1-2t)^3} & 0^-<t\leq\frac{1-\sqrt{r}}{2} \\ 0 & \text{otherwise}\end{cases}$$

for some $0<r\leq 1$, Z_T is a binary random variable with $\Pr(Z_T=1|T=t)=t$, and (Z_T,T) is statistically independent of X. It can be easily verified that $\Pr(\alpha_Y=t|T=t)=\Pr(\alpha_Y=1-t|T=t)=1/2$.

By [8, Theorem 1], the optimal quantizer partitions the interval $[0,1]$ into M subintervals $\mathcal{I}_i=[\gamma_{i-1},\gamma_i)$ for $i=1,\dots,M-1$ and $\mathcal{I}_M=[\gamma_{M-1},\gamma_M]$, where $0=\gamma_0<\gamma_1<\cdots<\gamma_M=1$, and outputs $f(y)=i$ iff $\alpha_y\in\mathcal{I}_i$. We therefore have

$$\begin{aligned} I(X;f(Y)) &= \sum_{i=1}^{M}\Pr(\alpha_Y\in\mathcal{I}_i)d\left(\mathbb{E}[\alpha_Y|\alpha_Y\in\mathcal{I}_i]\,\|\,\frac{1}{2}\right) \\ &\leq M\max_{0\leq a<b\leq 1}\Pr(a\leq\alpha_Y\leq b)d\left(\mathbb{E}[\alpha_Y|a\leq\alpha_Y\leq b]\,\|\,\frac{1}{2}\right). \end{aligned}$$

By the symmetry of the random variable α_Y around $1/2$, we can restrict the optimization to $a<1/2$ and $a<b\leq 1$. Let $\underline{b}=\min\{b,1-b\}$ and $\bar{b}=\max\{b,1-b\}$ and define the two intervals $\mathcal{T}_0=[a,\underline{b})$, $\mathcal{T}_1=[\underline{b},\bar{b}]$. By the convexity of KL divergence we have that

$$\begin{aligned} & d\left(\mathbb{E}[\alpha_Y|a\leq\alpha_Y\leq b]\,\|\,\frac{1}{2}\right) \\ &\leq \sum_{i=0}^{1}\Pr(\alpha_Y\in\mathcal{T}_i|a\leq\alpha_Y\leq b)d\left(\mathbb{E}[\alpha_Y|\alpha_Y\in\mathcal{T}_i]\,\|\,\frac{1}{2}\right) \\ &= \Pr(\alpha_Y\in\mathcal{T}_0|a\leq\alpha_Y\leq b)d\left(\mathbb{E}[\alpha_Y|a\leq\alpha_Y\leq\underline{b}]\,\|\,\frac{1}{2}\right), \end{aligned}$$

where we have again used the symmetry of the random variable α_Y in the last equation. We have therefore obtained

$$\begin{aligned} & I(X;f(Y)) \\ &\leq M\max_{0\leq a\leq b\leq\frac{1}{2}}\Pr(a\leq\alpha_Y\leq b)d\left(\mathbb{E}[\alpha_Y|a\leq\alpha_Y\leq b]\,\|\,\frac{1}{2}\right) \\ &= \frac{M}{2}\max_{0\leq a\leq b\leq\frac{1}{2}}\Pr(a\leq T\leq b)d\left(\mathbb{E}[T|a\leq T\leq b]\,\|\,\frac{1}{2}\right). \\ &= \frac{M}{2}\max_{0\leq b\leq\frac{1}{2}}\Pr(0\leq T\leq b)d\left(\mathbb{E}[T|0\leq T\leq b]\,\|\,\frac{1}{2}\right) \end{aligned}$$

where the last equality follows since both terms are individually maximized by $a=0$. It can be verified that for any $0\leq\rho\leq\frac{1-\sqrt{r}}{2}$

$$\int_0^{\rho}tf_T(t)dt=\frac{2r\rho^2}{(1-2\rho)^2};\ \Pr(0\leq T\leq\rho)=\frac{r}{(1-2\rho)^2},$$

and therefore $\mathbb{E}[T|0\leq T\leq b]=2b^2$, and we have that for any M-level quantizer

$$I(X;f(Y))\leq\frac{M}{2}\cdot\max_{0\leq b\leq\frac{1-\sqrt{r}}{2}}r\cdot\frac{1-h(2b^2)}{(1-2b)^2}\leq M\cdot\log(e)r,$$

where the last inequality follows by noting that the function $\frac{1-h(2b^2)}{(1-2b)^2}$ is monotone increasing in $0<b<1/2$, and taking the limit as $b\to 1/2$. It remains to relate r and $I(X;Y)$. Recalling that $h(\frac{1}{2}-p)\leq 1-2\log(e)p^2$, we have

$$\begin{aligned} I(X;Y)=1-\mathbb{E}h(T) &\geq 2\log(e)\mathbb{E}\left(\frac{1}{2}-T\right)^2 \\ &= 2\log(e)\frac{r}{4}\ln\left(\frac{e}{r}\right) \\ &= \frac{e\log(e)}{2}\frac{r}{e}\ln\left(\frac{e}{r}\right). \end{aligned}$$

Applying Proposition 6, we have

$$r\leq eg^{-1}\left(\frac{2I(X;Y)}{e\log(e)}\right)\leq\frac{2I(X;Y)}{\log(e)}\frac{1}{\ln\left(\frac{e\log(e)}{2I(X;Y)}\right)}$$

which gives

$$I(X;f(Y))\leq 2M\frac{I(X;Y)}{\ln\left(\frac{e\log(e)}{2I(X;Y)}\right)},$$

for any M-level function f. ■

Acknowledgment

The authors are grateful to Emre Telatar for pointing out the relation to the Knapsack problem, and to Guy Bresler, Meir Feder, Robert Gray, Pablo Piantanida, and Kevin Phillips for valuable discussions.

References

[1] N. Tishby, F. C. Pereira, and W. Bialek, "The information bottleneck method," in *37th Annual Allerton Conference on Communications, Control, and Computing*, Monticello, IL, pp. 368–377.

[2] A. J. Viterbi and J. K. Omura, *Principles of digital communication and coding*. Courier Corporation, 2013.

[3] T. Koch and A. Lapidoth, "At low snr, asymmetric quantizers are better," *IEEE Transactions on Information Theory*, vol. 59, no. 9, pp. 5421–5445, Sept 2013.

[4] R. Pedarsani, S. H. Hassani, I. Tal, and E. Telatar, "On the construction of polar codes," in *2011 IEEE International Symposium on Information Theory Proceedings*, July 2011, pp. 11–15.

[5] I. Tal, A. Sharov, and A. Vardy, "Constructing polar codes for non-binary alphabets and macs," in *2012 IEEE International Symposium on Information Theory Proceedings*, July 2012, pp. 2132–2136.

[6] A. Kartowsky and I. Tal, "Greedy-merge degrading has optimal power-law," *arXiv preprint arXiv:1701.02119*, 2017.

[7] B. M. Kurkoski and H. Yagi, "Quantization of binary-input discrete memoryless channels," *IEEE Transactions on Information Theory*, vol. 60, no. 8, pp. 4544–4552, Aug 2014.

[8] D. Burshtein, V. D. Pietra, D. Kanevsky, and A. Nadas, "Minimum impurity partitions," *The Annals of Statistics*, vol. 20, no. 3, pp. 1637–1646, 1992.

[9] H. Kellerer, U. Pferschy, and D. Pisinger, *Introduction to NP-Completeness of Knapsack Problems*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2004, pp. 483–493.