CHANNEL-OPTIMIZED STRATEGIC QUANTIZER DESIGN

Anonymous

ABSTRACT

We consider the design problem of a strategic quantizer over a noisy channel, extending the classical work on channel-optimized quantization to strategic settings. Building on the recent work on strategic quantization over noiseless channels, we employ a random channel index assignment mapping, as done in prior work on classical channel-optimized quantizer design literature, combined with a dynamic programming approach to optimize quantization boundaries. Our analysis and numerical results demonstrate several interesting aspects of channel optimized strategic quantization which do not appear in its classical (nonstrategic) counterpart.

Index Terms— Quantization, joint source-channel coding, game theory, dynamic programming

1. INTRODUCTION

This paper is concerned with the quantizer design problem for the setting where two agents (the encoder and the decoder) with misaligned objectives communicate over a noisy channel. The classical (non-strategic) counterparts of this problem have been investigated thoroughly in the literature, see e.g., [1, 2]. We here carry out the analysis to strategic communication cases, see e.g., [3–5] where the encoder and the decoder have different objectives, as opposed to the classical communication paradigm where the encoder and the decoder form a team with identical objectives.

Building on the recent work on strategic quantizer design over a perfect (noiseless) communication channel [5], and inspired by the prior literature on classical (nonstrategic) channel quantization via dynamic programming [2], we analyze and design the optimal channel quantizer for strategic settings, used in conjunction with random index assignment. Our main computational design tool is dynamic programming. We note that while dynamic programming has been utilized for channel-optimized quantizer design problems to avoid poor local minima issues common in nonconvex optimization [2], the strategic aspect of the problem makes the DP based solution the only viable solution approach as discussed in the recent work [5].

2. PROBLEM FORMULATION

Consider the following quantization problem: an encoder observes a realization of the source $X \in \mathcal{X}$ with a probability distribution μ and maps it to a message $Z \in \mathcal{Z}$, where $\mathcal{Z}$ is a set of discrete messages with a cardinality constraint $|\mathcal{Z}| \leq M$ using an injective mapping $Q : \mathcal{X} \to \mathcal{Z}$. An index mapping $\pi : [1 : M] \to [1 : M]$ is chosen uniformly at random and is applied to the message Z. The message $\pi(Z)$ is transmitted over a noisy channel with transition probability matrix $p(z_j|z_i)$. After receiving the message Z, the decoder applies a mapping $\phi : \mathcal{Z} \to \mathcal{Y}$ on the message Z and takes an action $Y = \phi(Z)$. The encoder and the decoder minimize their respective objectives $D_E = \mathbb{E}_\pi\{\mathbb{E}\{\eta_E(X, Y)|\pi\}\}$ and $D_D = \mathbb{E}_\pi\{\mathbb{E}\{\eta_D(X, Y)|\pi\}\}$, which are misaligned ($\eta_E \neq \eta_D$). The encoder designs Q *ex-ante*, i.e., without the knowledge of the realization of X, using only the objectives η_E and η_D, the statistics of the source $\mu(\cdot)$, and the channel parameters (transition probability matrix $p(z_j|z_i)$). The objectives (η_E and η_D), the shared prior (μ), the index assignment (π), the channel transition probability matrix ($p(z_j|z_i)$), and the mapping (Q) are known to the encoder and the decoder. The problem is to design Q for the equilibrium, i.e., the encoder minimizes its distortion if used with a corresponding decoder that minimizes its own distotion. The problem statement is presented below.

3. MAIN RESULTS

3.1. Analysis

Let X take values from the source alphabet $\mathcal{X} \in [a, b]$. The set $\mathcal{X}$ is divided into mutually exclusive and exhaustive sets $\mathcal{V}_1, \mathcal{V}_2, \ldots, \mathcal{V}_M$. The message

$$z_j = \pi(Q(x)), \quad x \in \mathcal{V}_i$$

where $Q(x) = z_i \,\forall x \in \mathcal{V}_i$, π is a bijective index mapping $\pi : \{1, \ldots, M\} \to \{1, \ldots, M\}$, is transmitted over a noisy channel and received as z'_j with probability $p(z'_j|z_j)$. The decoder receives the message and takes the action

$$y = \phi(z'_j).$$

We make the following “monotonicity” assumption.

Assumption 1. *$\mathcal{V}_m$ is convex for all $m \in [1 : M]$.*

Problem. *Using random index assignment for a given noisy channel with rate R and bit error rate p_{err}, purely discrete scalar source $x_0, \ldots, x_{N-1}$ points with a given probability mass $P(0), \ldots, P(N-1)$, find the decision boundaries $\mathbf{q} = [q_0, q_1, \ldots, q_M]$ and actions $\mathbf{y}(\mathbf{q}) = [y_1, \ldots, y_M]$ as a function of boundaries that satisfy:*

$$\mathbf{q}^* = \arg\min_{\mathbf{q}} \sum_{m=1}^{M} \mathbb{E}_\pi\{\mathbb{E}\{\eta_E(x, y_m(\mathbf{q}))|\pi, x \in [x_{q_{m-1}}, x_{q_m})\}\},$$

where actions $\mathbf{y}(\mathbf{q})$ are $y_m^*(\mathbf{q}) = \arg\min_{y_m \in \mathcal{Y}} \mathbb{E}_\pi\{\mathbb{E}\left\{\eta_D(x, y_m)|\pi, x \in [x_{q_{m-1}}, x_{q_m})\right\}\} \quad \forall m \in [1:M]$, *and the rate satisfies* $\log M \leq R$.

Under assumption 1, $\mathcal{V}_m$ is an interval since X is a scalar, i.e.,

$$\mathcal{V}_m = [x_{q_{m-1}}, x_{q_m})$$

where $0 = q_0 < q_1 < \ldots < q_M = N$. The encoder chooses the boundary indices $\mathbf{q} = [q_0, q_1, \ldots, q_M]$. The decoder determines its actions $\mathbf{y} = [y_1, \ldots, y_M]$ as the best response to $\mathbf{q}$ to minimize its cost $D_D = \mathbb{E}_\pi\{\mathbb{E}\{\eta_D(X, Y)|\pi\}\}$ for $m \in [1:M]$ as follows

$$y_m^* = \arg\min_{y_m \in \mathcal{Y}} \sum_{m=1}^{M} \mathbb{E}_\pi\{\mathbb{E}\{\eta_E(x, y_m)|\pi, x \in \mathcal{V}_m\}\}.$$

The average symbol error probability of the channel is

$$p_{err} = \frac{1}{M} \sum_{i=1}^{M} \sum_{j=1, j\neq i}^{M} p(z_j|z_i).$$

We limit the distortion (cost) functions to be polynomials[1], i.e.,

$$\eta_s(x, y) = \sum_{k \in \mathcal{K}} s_k x^{k_1} y^{k_2}$$

where $\mathcal{K} = \{(k_1, k_2)|k_1, k_2 \in \mathbb{Z}\}, s_k \in \mathbb{R}, s \in \{E, D\}$.

The integrals expressed throughout this paper are defined over the set $\mathcal{V}_i$, which is omitted for brevity. The probability that the receiver receives the noisy message $\hat{z} = z_j$ if z_i was transmitted using $\pi(Q(x)) = z_i$ is given by $p(z_j|z_i)$, the channel transition probability.

The end-to-end distortion given an index assignment π is

$$\mathbb{E}\{\eta_s|\pi\} = \sum_{i=1}^{M} \int \sum_{j=1}^{M} \sum_{k \in \mathcal{K}} s_k x^{k_1} y^{k_2} p(z_j|z_i) \mathrm{d}\mu(x).$$

[1]The choice polynomials does not introduce any loss of generality here since by Stone–Weierstrass theorem [6] any continuous function can be uniformly approximated by a sequence of multivariate polynomials.

The average distortion over all possible index assignments is

$$\begin{aligned} \overline{D^s} &= \mathbb{E}_\pi\{\mathbb{E}\{\eta_s|\pi\}\} \\ &= \sum_{i=1}^{M} \int \sum_{j=1}^{M} \sum_{k \in \mathcal{K}} s_k x^{k_1} y^{k_2} \mathbb{E}_\pi(p(z_j|z_i)) \mathrm{d}\mu(x) \\ &= \sum_{k} \left(I_{k, j\neq i} + I_{k, j=i}\right) \end{aligned}$$

where $I_{k,j\neq i}$ and $I_{k,j=i}$ are defined as follows:

$$I_{k,j\neq i} = s_k \sum_{i=1}^{M} \int \sum_{\substack{j=1 \\ j\neq i}}^{M} x^{k_1} y^{k_2} \mathbb{E}_\pi(p(z_j|z_i)) \mathrm{d}\mu(x),$$

$$I_{k,j=i} = s_k \sum_{i=1}^{M} \int x^{k_1} y^{k_2} \mathbb{E}_\pi(p(z_i|z_i)) \mathrm{d}\mu(x).$$

We next express these terms in a way that can be approached via dynamic programming:

$$\begin{aligned} I_{k,j\neq i} &= s_k \sum_{i=1}^{M} \int \sum_{j=1, j\neq i}^{M} x^{k_1} y_j^{k_2} \mathbb{E}_\pi(p(z_j|z_i)) \mathrm{d}\mu(x) \\ &= \frac{p_{err}}{M-1} s_k \sum_{i=1}^{M} \int x^{k_1} \sum_{j=1, j\neq i}^{M} y_j^{k_2} \mathrm{d}\mu(x) \\ &= \frac{p_{err}}{M-1} s_k \sum_{i=1}^{M} \int x^{k_1} \left(\sum_{j=1}^{M} y_j^{k_2} - y_i^{k_2}\right) \mathrm{d}\mu(x), \end{aligned}$$

$$\begin{aligned} I_{k,j=i} &= s_k \sum_{i=1}^{M} \int x^{k_1} y_i^{k_2} \mathbb{E}_\pi(p(z_i|z_i)) \mathrm{d}\mu(x) \\ &= (1 - p_{err}) s_k \sum_{i=1}^{M} y_i^{k_2} \int x^{k_1} \mathrm{d}\mu(x). \end{aligned}$$

Let $c_1 = \frac{p_{err}}{M-1}$, $c_2 = 1 - Mc_1$. We assume $0 < p_{err} < \frac{M-1}{M}$ so that $c_1, c_2 > 0$. Then, the average distortion and the optimum decoder reconstruction for $i \in [1:M]$ are given by

$$\overline{D^s} = \sum_{k \in \mathcal{K}} s_k \sum_{i=1}^{M} y_i^{k_2} \left(c_1 \mathbb{E}\{x^{k_1}\} + c_2 \int x^{k_1} \mathrm{d}\mu(x)\right), \quad (1)$$

$$y_i = \arg\min_{y\in\mathcal{Y}} \sum_{k\in\mathcal{K}} s_k y^{k_2}\left(c_1\mathbb{E}\{x^{k_1}\} + c_2\int x^{k_1}\mathrm{d}\mu(x)\right). \quad (2)$$

3.2. Dynamic Programming Algorithm

While our analysis holds for general source alphabets, for the dynamic programming algorithm, the source has to be purely discrete. In case it is not, we approximate it by uniformly quantizing $[a, b]$ to N points to obtain an ordered set of N points, i.e.,

$$\mathcal{X} = \{a + (1 + 2t)\delta\}, \quad t = 0, \ldots, N-1$$

where $\delta = \frac{b-a}{2N}$ with the probability mass function

$$P(t) = \int_{x_t-\delta}^{x_t+\delta} d\mu, \quad t \in [0 : N-1].$$

Let $\mathcal{X}$ be an ordered (ascending) set in $[a, b]$,

$$\mathcal{X} = \{x_t\}, \quad t = 0, \ldots, N-1,$$

where $a = x_0 < x_1 < \ldots < x_{N-1} = b$ with a probability mass function P. We carry out the analysis in discrete setting, which yields that the integrals in equations (1) and (2) transform into summations over $x_t \in \mathcal{V}_i$. We define the following expressions:

1. The encoder and decoder costs for source interval $[x_\alpha, x_\beta)$ for a given action y for $s \in \{E, D\}$:

$$C_s(\alpha, \beta, y) = \sum_{k\in\mathcal{K}} s_k y^{k_2}\left(c_1\mathbb{E}\{x^{k_1}\} + c_2\sum_{t=\alpha}^{\beta-1} x_t^{k_1} P(t)\right).$$

2. The decoder's optimal action for the interval $[x_\alpha, x_\beta)$:

$$\kappa(\alpha, \beta) = \arg\min_{y\in\mathcal{Y}} C_D(\alpha, \beta, y).$$

3. Costs for the source interval $[x_\alpha, x_\beta)$ in conjunction with the optimal action for $s \in \{E, D\}$:

$$\epsilon_s(\alpha, \beta) = C_s(\alpha, \beta, \kappa(\alpha, \beta)).$$

4. Equilibrium costs associated with the m level optimal strategic quantizer for $[x_0, x_n)$ for $s \in \{E, D\}$, where $\mathbf{q} = [q_0, \ldots, q_m], 0 = q_0 < \ldots < q_m = n$:

$$D_s(n, m) = \min_{\mathbf{q}} \sum_{j=1}^{m} \epsilon_s(q_{j-1}, q_j).$$

5. The set of all non-empty convex subsets of $[0 : N]$:

$$\mathcal{S} = \{[t_1, t_2) : t_1, t_2 \in [0 : N], t_1 < t_2\}.$$

We let $\mathbf{q}$ be the optimal $m-$level strategic quantizer applied to $[x_0, x_n)$. The encoder's cost associated with the optimal m level quantization of $[x_0, x_n)$ can be written as the sum of that for the $(m-1)$ level quantization of $[x_0, x_v), v < n$ and 1 level quantization of $[x_v, x_n)$. Let $h(n, m) = q_{m-1}$ be the $(m-1)^{th}$ decision level of $\mathbf{q}$. The backward induction yields the following set of equations for $m \in [2 : n]$:

$$D_E(n, m) = D_E(h(n, m), m-1) + \epsilon_E(h(n, m), n),$$
$$h(n, m) = \arg\min_{i\in[m-1:n-1]} \{D_E(i, m-1) + \epsilon_E(i, n)\}.$$

Dynamic programming requires a forward and a backward pass. During the forward pass, we compute and store $h(n, m)$ and $D_E(n, m)$ for each pair $(n, m), m \in [2 : M], n \in [m : N]$. To do this, we first compute $\epsilon_E(\alpha, \beta)$ for all $[\alpha, \beta) \in \mathcal{S}$. We set $D_E(n, 1) = \epsilon_E(0, n), n \in [1 : N]$. The values of $h[n, m]$ and $D_E(n, m)$ are recursively calculated starting from $m = 2, n \geq m$.

In the backward pass, we set $q_0^* = 0, q_M^* = N$ and compute $\{q_m^*\}$ recursively as

$$q_{m-1}^* = h(q_m^*, m), \quad m = M, \ldots, 2.$$

The optimal representative levels are given by

$$y_m^* = \kappa(q_{m-1}^*, q_m^*), \quad m = [1 : M].$$

3.3. Special Distortion Functions

If the distortion functions are of a specific form given below

$$\eta_E(x, y) = (f(x) - y)^2, \quad \eta_D(x, y) = (x - y)^2 \quad (3)$$

where $f(\cdot)$ is any (Borel) measurable function, then the analysis simplifies as follows: We first express the encoder's distortion without noise or random index assignment as:

$$D_Q^E = \sum_{i=1}^{M} \int (f(x) - y_i)^2 \mathrm{d}\mu(x).$$

The average encoder distortion over all possible index assignments is then

$$\overline{D^E} = \mathbb{E}_\pi\{\sum_{i=1}^{M} \int \sum_{j=1}^{M} (f(x) - y_j)^2 p(z_j|z_i)\mathrm{d}\mu(x)\}$$
$$\stackrel{(1)}{=} c_1 \sum_{i=1}^{M} \int \sum_{j=1}^{M} (y_i - y_j)^2 \mathrm{d}\mu(x)$$
$$+ c_1 \sum_{i=1}^{M} \int \sum_{j=1}^{M} 2(f(x) - y_i)(y_i - y_j)\mathrm{d}\mu(x) + D_Q^E$$
$$\stackrel{(2)}{=} c_1 \sum_{i=1}^{M} \left(y_i - f_1\right)^2 + M c_1 \mathbb{E}\{(f(x) - f_1)^2\} + c_2 D_Q^E$$

where we denote $\mathbb{E}\{f(x)\}$ and $\mathbb{E}\{f^2(x)\}$ as f_1 and f_2 respectively. Steps 1 and 2 above are obtained by expressing the term $(f(x) - y_j)^2$ as $(f(x) - y_i + y_i - y_j)^2$, and by using the following expansion of D_Q^E respectively.

$$
\begin{aligned}
D_Q^E &= \sum_{i=1}^{M} \int (f(x) - y_i)^2 \mathrm{d}\mu(x) \\
&= f_2 + \sum_{i=1}^{M} y_i^2 \int \mathrm{d}\mu(x) - 2\sum_{i=1}^{M} y_i \int f(x)\mathrm{d}\mu(x).
\end{aligned}
$$

The terms $C_E(\alpha, \beta, y)$ and $D_E(n, m)$ then simplify as

$$
C_E(\alpha, \beta, y) = c_1(y - f_1)^2 + \sum_{t=\alpha}^{\beta-1} c_2 P(t)(f(x_t) - y)^2,
$$

$$
D_E(n, m) = \min_{\mathbf{q}} \sum_{j=1}^{m} \epsilon_E(q_{j-1}, q_j) + Mc_1(f_2 - f_1^2).
$$

The actions associated with $\mathcal{V}_i$ are given by minimizing $J_i(y)$,

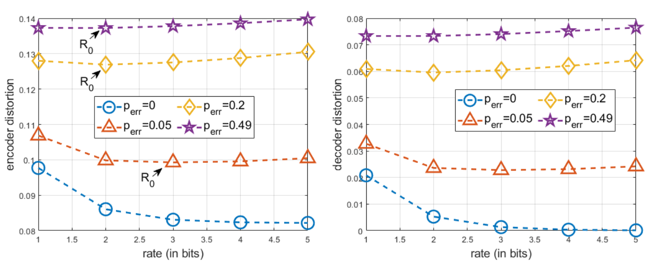
$$
J_i(y) = c_1(y - \mathbb{E}\{x\})^2 + c_2 \int (x - y)^2 \mathrm{d}\mu(x),
$$

$$
y_i = \arg\min_{y \in \mathcal{Y}} J_i(y).
$$

The first-order optimality condition, $\frac{\partial J_i}{\partial y} = 0$, yields

$$
y_i = \frac{c_1 \mathbb{E}\{x\} + c_2 \int x \mathrm{d}\mu(x)}{c_1 + c_2 \int d\mu(x)}.
$$

4. NUMERICAL RESULTS



(a) Encoder distortion (b) Decoder distortion

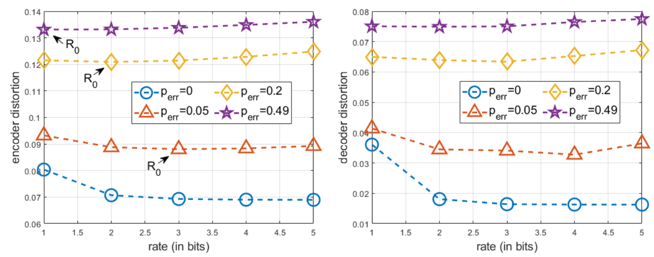
Fig. 1. $\eta_E = (x - 1.5y)^2, \eta_D = (x - y)^2$

We consider a source uniformly distributed on $[0, 1]$, i.e., $X \sim U(0, 1)$, and a binary symmetric channel with crossover probability $p_b \le \frac{1}{2}$, which yields

$$
p_{err} = 1 - (1 - p_b)^{\log M}.
$$

We take the decoding distortion measure as $\eta_D(x, y) = (x - y)^2$, and consider two different cases of encoder distortion, $\eta_E^1(x, y) = (x - 1.5y)^2$ and $\eta_E^2(x, y) = (x^3 - y)^2$.

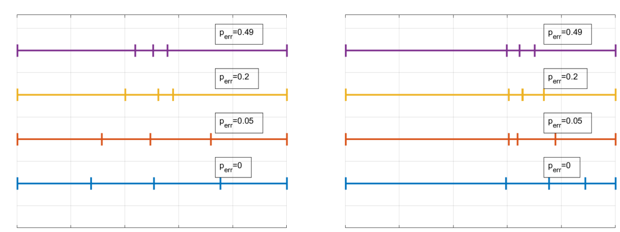
We plot the encoder and the decoder distortions associated with η_D in conjunction with η_E^1 and η_E^2 in Figures 1 and 2 respectively. One surprising aspect of these results is that the encoder distortion may be be increasing with rate at high resolution. This is due to the strategic aspect of the problem, i.e., at high resolution, the encoder is forced to be more revealing than its optimal choice. The impact of channel noise can be seen in the rate threshold, i.e., the cutoff rate, the smallest R_0 for which $R > R_0$ implies $D(R) \ge D(R_0)$. Numerical results shown in Figures 1 and 2, suggest that as p_{err} increases, the cutoff rate R_0 gets smaller. This observation indicates that in the high-rate regime, the optimal strategic encoder might choose not to utilize the channel rate fully.



(a) Encoder distortion (b) Decoder distortion

Fig. 2. $\eta_E = (x^3 - y)^2, \eta_D = (x - y)^2$

We plot the obtained quantizers in Figure 3. The numerical results depicted below suggest that the encoder is less revealing with increasing p_{err}. While due to space constraints, we only present the results associated with a uniform distribution here, results for a Gaussian source, as well as the codes that produced these results are available at [7].



(a) $\eta_E = (x - 1.5y)^2$ (b) $\eta_E = (x^3 - y)^2$

Fig. 3. Quantizers for $M = 4$ with $\eta_D = (x - y)^2$

5. CONCLUSIONS

In this paper, we extended the DP-based strategic quantizer design approach to the problem of strategic quantization with channel noise. Our analysis and the obtained numerical results have uncovered several aspects of the optimal strategic quantizers in this noisy channel setting which are not observed in its classical (nonstrategic) counterpart.

6. REFERENCES

[1] Xiang Yu, Haiquan Wang, and En-Hui Yang, "Design and analysis of optimal noisy channel quantization with random index assignment," *IEEE Transactions on Information Theory*, vol. 56, no. 11, pp. 5796–5804, 2010.

[2] Sorina Dumitrescu, "On the design of optimal noisy channel scalar quantizer with random index assignment," *IEEE Transactions on Information Theory*, vol. 62, no. 2, pp. 724–735, 2016.

[3] Emrah Akyol, Cédric Langbort, and Tamer Başar, "Information-theoretic approach to strategic communication as a hierarchical game," *Proceedings of the IEEE*, vol. 105, no. 2, pp. 205–218, 2016.

[4] Emir Kamenica and Matthew Gentzkow, "Bayesian persuasion," *American Economic Review*, vol. 101, no. 6, pp. 2590–2615, 2011.

[5] Anju Anand and Emrah Akyol, "Optimal strategic quantizer design via dynamic programming," in *Proceedings of the IEEE Data Compression Conference (DCC)*. IEEE, 2022, pp. 173–181.

[6] Halsey Lawrence Royden and Patrick Fitzpatrick, *Real analysis*, vol. 32, Macmillan New York, 1988.

[7] Anju Anand and Emrah Akyol, "Strategic quantization with noise matlab codes," `https://github.com/anjuanand98/dpnoise`, 2022, [Online; accessed 16-Oct-2022].