Strategic Quantization

Emrah Akyol
Electrical and Computer Engineering
Binghamton University, SUNY NY, USA
https://orcid.org/0000-0002-0663-1677
eakyol@binghamton.edu

Anju Anand
Electrical and Computer Engineering
Binghamton University, SUNY NY, USA
https://orcid.org/0000-0002-9378-5833
aanand6@binghamton.edu

***Abstract*—This paper is concerned with the quantization setting where the encoder and the decoder have misaligned objectives. While the unconstrained variation of this problem has been well-studied under the theme of information design problems in Economics, the problem becomes more appealing and relevant to engineering applications with a constraint on the cardinality of the message space. We first motivate the problem via a toy example demonstrating the intricacies of the strategic variation of the quantization problem, explicitly showing that the quantization resolution can change the nature of optimal encoding policy and the iterative optimization of the decoder and the encoder mappings may not converge to a local optimum solution, in sharp contrast with the classical quantization problem that admits the Lloyd-Max method as a local optimum solution. As a remedy, we develop a gradient-descent based solution. We analyze the poor local optimal optima issues associated with the optimization method and show that even for well-behaving sources, such as Gaussian, there are multiple local minima, depending on the distortion measures chosen, in sharp contrast with the classical quantization. We finally present numerical results obtained via the proposed algorithm that suggest their validity and demonstrate features of the strategic version of the quantization problem that differentiate it from its classical counterpart.**

I. Introduction

Consider the following problem: an encoder observes a realization of a source $X \in \mathcal{X}$ with a probability distribution μ and sends a message $Z \in \mathcal{Z}$ using a non-injective mapping $Q : \mathcal{X} \to \mathcal{Z}$, with $|\mathcal{Z}| \leq M$. After receiving the message Z, the decoder takes action $Y \in \mathcal{Y}$. The costs that the encoder and the decoder minimize are $D_E \triangleq \eta_E(X, Y)$ and $D_D \triangleq \mathbb{E}\{\eta_D(X, Y)\}$, with $\eta_E \neq \eta_D$ (misaligned objectives). The encoder designs Q *ex-ante*, i.e., without the knowledge of the realization of X, using only the functions η_E and η_D, and the statistics of the source, $\mu(\cdot)$. The functions (η_E and η_D), the shared prior (μ), and the mapping (Q) are known to the encoder and the decoder. The problem is to design Q. We call this setup *strategic quantization*, which is the focus of this paper.

The setting without the quantization aspect (in a practical sense, if M is asymptotically large) is known in the Economics literature as the information design, or the Bayesian persuasion problem [1], [2]. These problems analyze how a communication system designer (sender) can use the information to influence the action taken by a receiver. This framework has proven beneficial in analyzing a variety of real-life applications, such as the design of transcripts when schools compete to improve their students' job prospects [3] and voter mobilization and gerrymandering [4], as well as various engineering applications, including in modeling misinformation spread over social networks [5] and privacy-constrained information processing [6], and many more [7]. For an excellent survey of the related literature in Economics, see [8], [9].

The following scenario constitutes an engineering example of the setting studied here. Consider two smart cars from different brands, such as Tesla and Honda, where Tesla (as the decoder) asks Honda (the encoder), for information about the traffic ahead to decide whether it should change its route. However, Tesla and Honda have different goals that are known to both. Tesla wants to accurately estimate the true traffic condition, but Honda wants to make Tesla do something else, like changing its route. In this scenario, Honda will not necessarily fully reveal the traffic information, because its goal is not the same as Tesla's. To make Tesla use the transmitted information, Honda has to make sure that Tesla benefits from following its information, that is, the Tesla's cost in using Honda's information is smaller than that in ignoring it. Assuming a fixed-rate communication channel between them, how would these cars communicate? Our analysis in this paper provides a theoretical framework for such class of communication problems.

The strategic quantization problem, as described above, was discussed in a few contemporary economics and computer science studies. In [10], authors analyze the problem via a computation lens and report approximate results on this problem, relating to another problem they solved conclusively. In one of their main results, the algorithmic complexity of finding the optimum strategic quantizer was shown to be NP-hard. In a recent working paper, Aybaş and Türkel [11] analyzed this problem using the methods in [2] and provided several theoretical properties of strategic quantization. A byproduct of their analysis yields a constructive method for deriving optimal quantizers based on a search over possible posterior distributions over their feasible set. Our objective here is to leverage the rich collection of results in quantization theory, e.g., the comprehensive survey of results by Gray and Neuhoff [12], to study the same problem via the engineering lens.

We note in passing that quantizers also arise as equilibrium strategies endogenously, i.e., without an external constraint, in a related but distinctly different class of signaling games, namely the cheap talk [13], [14] In [13], the encoder chooses the mapping from the realization of the source X to message Z *after* observing it, *ex-post*, as different source realizations indicate optimality of different mappings for the encoder. The encoder's lack of commitment power in the cheap talk setting makes the notion of equilibrium a Nash equilibrium since both agents form a strategy that is the best response to each other's mapping. However, in our strategic quantization problem (and the information design problems in general as in [1], [2]), the encoder designs Q *ex-ante*, *before* seeing the source realization, and committed to the designed Q afterward. This commitment is known to the decoder and establishes a form of trust between the sender and the receiver, resulting in possibly higher payoffs for both agents. This difference also manifests itself in the notion of equilibrium we are seeking here since the encoder does not necessarily form the best response to the decoder due to its commitment to Q[1].

II. Model

The set $\mathcal{X}$ is divided into mutually exclusive and exhaustive sets $\mathcal{V}_1, \mathcal{V}_2, \dots, \mathcal{V}_M$. The action of the decoder y_m can be written as a quantization operator

$$y_m = Q(x) \quad \forall x \in \mathcal{V}_m$$

for all $m \in [1:M]$. Throughout this paper, we make the following "monotonicity" assumption on the sets $\{\mathcal{V}_m\}$.

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

Remark 1: Assumption 1 is the first of the two regularity conditions commonly employed in the classical quantization literature, cf. [16]. Note that the second regularity condition, $y_m \in \mathcal{V}_m$, is not included in Assumption 1.

In the Economics parlance, Assumption 1 is referred as the "monotonicity" condition. In [17], [18] sufficient conditions on η_E and η_D for the monotonicity of optimal encoder strategies are characterized within the unconstrained (without quantization) variation of the same problem. We note that here we have the quantization constraint in the problem formulation as an exogenous constraint on the message set, hence it is not clear apriori whether the results [17], [18] would be applicable here. Under Assumption 1, $\mathcal{V}_m$ is an interval, i.e.,

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

where $a = x_0 < x_1 < \ldots < x_M = b$. The encoder chooses the boundary indices $\mathbf{q} = [x_0, x_1, \dots, x_M]$. The decoder determines its actions $\mathbf{y} = [y_1, \dots, y_M]$ as a best response to $\mathbf{q}$ to minimize its cost $D_D = \mathbb{E}\{\eta_D(x, y)\}$ as follows

$$y_m^* = \arg\min_{y_m \in \mathcal{Y}} \mathbb{E}\{\eta_D(x, y_m) | x \in \mathcal{V}_m\} \quad \forall m \in [1:M]$$

Hence, the decoder chooses the actions $\{y_m\}$ knowing the set of decision sets $\{\mathcal{V}_m\}$. The encoder computes what the decoder would choose as $\mathbf{y}$ given $\{\mathcal{V}_m\}$, and hence optimizes its own cost $\mathbb{E}\{\eta_E(x, y)\}$ over the choice of $\{\mathcal{V}_m\}$ accordingly:

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

or due to Assumption 1 equivalently over the choice of $\mathbf{q}$:

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

Remark 2: A key consideration here is that the encoder is committed to its choice of $\mathbf{q}$, it cannot determine $\mathbf{q}$ as the best response to $\mathbf{y}$. Hence, while the decoder can optimize its action $\mathbf{y}$ as the best response to $\mathbf{q}$, the encoder cannot choose $\mathbf{q}$ as the best response to $\mathbf{y}$, but that to a function of itself, i.e., the best response to $\mathbf{y}(\mathbf{q})$. This aspect of the problem introduces a hierarchy in the game play (the encoder plays first, and the decoder responds, which is referred to as the "Stackelberg equilibrium" in the computer science and control literature, and more formally constitutes an instance of subgame perfect Bayesian Nash equilibrium) and naturally is not a Nash equilibrium since $\mathbf{q}$ may not be the best response to $\mathbf{y}$. In cheap talk [13], Nash equilibria are sought after and the equilibria achieving strategies happen to be non-injective mappings, i.e., quantizers, without an exogenous rate constraint. It is essential to note the substantial difference between the problem formulation in this paper and the cheap talk literature [13].

III. Interesting Aspects of Strategic Quantization

In this subsection, we provide simple numerical examples to demonstrate a few intricacies of the strategic quantization problem that differentiate it from its classical analogue.

A. To reveal or not to reveal... or partially reveal?

As mentioned earlier, the classical quantization problem is a team problem, i.e., the encoder and the decoder share identical objectives. In strategic quantization, this is no longer the case. Hence, depending on the misalignment between η_E and η_D, the encoder might choose not to send any information to the decoder (e.g., if they are too misaligned), which we refer to as "non-revealing" policy following the convention in the Economics literature. Alternatively, the strategic quantizer problem might simplify to classical quantization with a common distortion measure η_D, i.e., the encoder cannot utilize its information design advantage to persuade the decoder to take a specific action (referred to here as "fully-revealing" encoder policy). Finally, the encoder might choose to employ a quantizer that is not identical to the classical one, i.e., "partially-revealing" policy. It is relatively straightforward to provide an example for the first case; consider, e.g., $\eta_E(x, y) = -\eta_D(x, y)$, $\forall x, y$ which makes the problem a zero-sum game, hence, the optimal strategy for the encoder is not to send any information (see also the analysis below). We demonstrate the latter cases via a simple numerical example.

[1]These issues are well understood in the Economics literature, see, e.g., [15] for an excellent survey. However, we emphasize them here for a reader with an engineering background; see [7] for a detailed discussion through the engineering lens.

Consider a setting where the decoder wants to reconstruct X as closely as possible, while the encoder wants the decoder's construction to be as close as possible to $\frac{X+\alpha}{\beta}$, both in the MSE sense. Can the encoder "persuade" the decoder by carefully designing quantizer intervals $\mathcal{V}_m^*$? We prove the following result, which is an extension of a result presented in [18], in the Appendix:

Theorem 2: For $\eta_E(x,y) = (x+\alpha-\beta y)^2$ and $\eta_D(x,y) = (x-y)^2$, the optimal strategic quantizer Q is:

$$Q(x) = \left\{ \begin{array}{ll} \arg\min \mathbb{E}\{(X-Q(X))^2\}, & \text{for } 0<\beta<2 \\ arbitrary, & \text{for } \beta = 0,2 \\ constant, & \text{otherwise} \end{array} \right\}.$$

Note that the first case corresponds to the fully-revealing behavior, while the second one corresponds to encoder distortion remaining constant for all quantizers, and the third is non-revealing.

Remark 3: This theorem highlights the difference between the problem formulation here and that of [13] which considers identical cost functions with $\beta = 1$. Our result indicates that there is no game problem if these cost functions are used in the Stackelberg setting, while in the Nash equilibrium setting [13] , i.e., without the commitment power, the game nature of the problem is well understood.

B. Quantizer resolution is binding

We next focus on the question: can the quantization constraint change the nature of optimal encoding policy in strategic communication? For example, is there a case where for $M = 2$ the encoder is non-revealing but for $M = 3$ the encoder prefers to send a message? The answer is, perhaps surprisingly, affirmative.

Consider a continuous source $X \sim U[-1,1]$ with the distortion functions $\eta_D(x,y) = (x-y)^2$ and

$$\eta_E(x,y) = \begin{cases} (x^3-y)^2, & xy \geq 0 \\ \to \infty, & \text{otherwise.} \end{cases}$$

A fully non-revealing policy, i.e., the case of $R = 0$ ($M = 1$) yields $D_E(0) = \int_{-1}^{1} (t^3)^2 \frac{1}{2}\, dt = 1/7$.

We next consider $M = 2$. From symmetry, the optimal encoding policy is simply setting the boundary at $q_1 = 0$. However, this yields $y_1 = -y_2 = -1/2$ and $D_E(1) = 1/7$, which is identical to $D_E(0)$. Hence, the optimal strategic quantizer for $M = 2$ does not send any information to the decoder, i.e., non-revealing.

We finally consider the case of $M = 3$. Similar to the previous case, we parameterize $\mathcal{V}_m^*$ as $[-1,q], (q,0], (0,1]$ (parameterizing $\mathcal{V}_m^*$ as $[-1,0], (0,q], (q,1]$ results in a solution violating constraints) and express $\{y_m\}$ as a function of q, which yields

$$y_1 = \frac{-1+q}{2}, \quad y_2 = \frac{q}{2}, \quad y_3 = \frac{1}{2}.$$

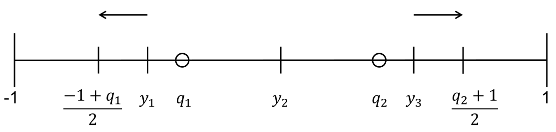


Fig. 1: Movement of the quantization boundaries through Lloyd-Max iterations in the running example.

Substituting again $\{y_m\}$ in $\eta_E(x,y)$:

$$J(q) = \frac{1}{2}\bigg(\int_{-1}^{q} (t^3-y_1)^2\, dt + \int_{q}^{0} (t^3-y_2)^2\, dt + \int_{0}^{1} (t^3-y_3)^2\, dt \bigg).$$

Enforcing the KKT conditions similar to (3), we obtain $q = -0.8090, 0.3090, 0.5$. Since $q \leq 0$, the possible solution is $q = -0.8090$. The encoder is partially revealing here: in the non-strategic case ($\eta_E = \eta_D = (x-y)^2$), the quantizer for $X \sim U[-1,1]$ would have been a uniform quantizer. However, here, the encoder uses a different quantizer, that is, the encoder does not reveal information exactly as the decoder would want.

We note here that $D_E(3) = 0.1146 < 1/7 = D_E(1) = D_E(0)$. Hence, at $M = 2$, the strategic quantizer does not communicate any information while, at $M = 3$, in sharp contrast to $M = 2$, uses the quantization channel fully to send three messages, demonstrating that the quantization constraint can change the nature of the optimum encoder policy. Moreover, it shows that the operational rate-distortion function $D_E(R)$ here is not a strictly decreasing function of rate R, since $D_E(1) = D_E(0)$, unlike its classical counterpart.

C. Failure of "Strategic Lloyd-Max"

In this subsection, we investigate whether a simple strategic variation of the Lloyd-Max approach would yield a locally optimal quantizer, as in the case of classical quantization. Consider a continuous source $X \sim U[-1,1]$ quantized with $M = 3$ messages where $\eta_E(x,y) = (x^3-y)^2$ and $\eta_D(x,y) = (x-y)^2$.

We initialize $\{y_m\}$ arbitrarily and find q_1 and q_2 that minimize $D_E = \mathbb{E}\{(x^3-y)^2\}$, as

$$q_1 = \left(\frac{y_1+y_2}{2}\right)^{\frac{1}{3}}, q_2 = \left(\frac{y_2+y_3}{2}\right)^{\frac{1}{3}}. \tag{1}$$

We then find the decoder actions $\{y_m\}$

$$y_1 = \frac{\int_{-1}^{q_1} \frac{1}{2} t\ dt}{\int_{-1}^{q_1} \frac{1}{2}\ dt} = \frac{-1+q_1}{2}, y_2 = \frac{\int_{q_1}^{q_2} \frac{1}{2} t\ dt}{\int_{q_1}^{q_2} \frac{1}{2}\ dt} = \frac{q_1+q_2}{2},$$

$$y_3 = \frac{\int_{q_2}^{1} \frac{1}{2} t\ dt}{\int_{q_2}^{1} \frac{1}{2}\ dt} = \frac{q_2+1}{2}, \tag{2}$$

and iterate between (1) and (2) until convergence. We note that during these iterations, q_1 and q_2 move towards -1 and 1, respectively, i.e., the boundaries move towards the endpoints of the interval taken for quantization with each iteration as demonstrated in Figure 1, hence the iterations converge to

$M = 1$ solution which is non-revealing with $D_E = 1/7$. We next examine whether this solution is a local optimum.

Any admissible perturbation of with some $1 > \epsilon > 0$ of $q_1 = -q_2 = -1$ would result in a $M = 3$ level quantizer with decision boundaries $q_1 = -1 + \epsilon$, $q_2 = 1 - \epsilon$ with the corresponding decoder actions $y_1 = -1 + \frac{\epsilon}{2}$, $y_2 = 0$, $y_3 = 1 - \frac{\epsilon}{2}$, yielding $D_E = 1/7 - \epsilon(1-\frac{\epsilon}{2})^2(1-\epsilon)^2$ which is smaller than that of the non-revealing solution ($D_E = 1/7$), hence this is not a locally optimal solution. This observation indicates that the straightforward enforcement of optimality conditions may not yield a locally optimal solution, which contrasts sharply with the case in classical quantization. In other words, unlike its classical counterpart, a trivial extension of the Lloyd-Max algorithm adopted for strategic settings may not converge to a locally optimum solution.

D. Multiple Local Minima

A well-known fact in classical quantization literature is that for log-concave scalar sources, the local minima coincides with the global one, hence Lloyd-Max is guaranteed to converge to the globally optimal solution. Hence, a natural follow-up question here pertains to whether a similar result holds in the strategic realm. We deduce the answer to this question via a numerical counter-example: Consider a Uniform scalar source $X \sim U[a, b]$ with $\eta_E = (x^3 - y)^2, \eta_D = (x - y)^2$. The decoder actions y_1, y_2,

$$y_1 = \frac{a+q}{2}, \quad y_2 = \frac{q+b}{2}.$$

The cost function associated with two-level quantization of X ($\mathbf{q} = \begin{bmatrix} a & q & b \end{bmatrix}$) and it's derivative with respect to the quantization decision level q,

$$J(\mathbf{q}) = \int_a^q (x^3 - y_1)^2 \mathrm{d}\mu + \int_q^b (x^3 - y_2)^2 \mathrm{d}\mu,$$

$$\frac{\partial J}{\partial q} = q^3 - \frac{q}{2} + \frac{a+b}{4}(1 - (b^2 + a^2)).$$

For $a = -0.9$, $b = 1$, $q = -0.6859, 0.7265$ satisfy $\frac{\partial J}{\partial q} = 0$, as also demonstrated in Fig 2. Hence, the strategic problem can indeed have multiple local minima even if used in conjunction with log-concave sources.

IV. Proposed Algorithm

Having shown that a strategic variation of the Lloyd-Max algorithm does not converge, we propose a gradient-descent based solution. Due to the Stackelberg equilibrium nature of the problem, the encoder (leader) decides the quantizer decision levels q first, and then the decoder (follower) determines the quantizer reconstruction levels as a function of $\mathbf{q}$, as $\mathbf{y}(\mathbf{q})$. So the gradient descent approach involves optimizing $\mathbf{q}$ along the direction of the gradients $\frac{\partial D_E}{\partial q_m}$.

Starting with an initial set of quantizers $\mathbf{q} = \mathbf{q}_0$, the reconstruction levels $\mathbf{y}(\mathbf{q})$, and the associated encoder distortion D_E are computed. Then, the following steps are executed until convergence:

1) The gradients $\{\frac{\partial D_E}{\partial x_m}\}$ are computed.

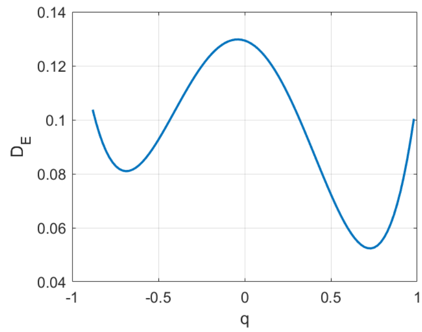


Fig. 2: Local minima of the cost surface of the encoder distortion for 2-level quantization, with decision boundary q, of $X \sim U[-0.9, 1]$ with $\eta_E = (x^3 - y)^2, \eta_D = (x - y)^2$.

2) The decision levels $\mathbf{q}$ are updated if the resulting quantizer adhered to the quantizer constraints and the distortion is non-increasing with the update: $x_m \triangleq x_m - \triangle \frac{\partial D_E}{\partial x_m}$, where $\triangle$ is the gradient descent parameter; else Δ is decreased repeatedly until the conditions are satisfied.
3) The reconstruction levels $\mathbf{y}(\mathbf{q})$ are found.
4) The encoder distortion D_E is computed.

Remark 4: The algorithm is guaranteed to converge to a locally optimal solution under the assumptions in this paper. We present below the derivation of gradients for an MMSE decoder $\eta_D = (x - y)^2$ with an encoder with distortion measure $\eta_E(x, y)$. The gradients of the encoder's distortion with respect to the quantizer decision levels are given as:

$$\begin{aligned}\frac{\partial D_E}{\partial x_m} &= \eta_E(x_m, y_m)\frac{\mathrm{d}\mu(x_m)}{\mathrm{d}x} - \eta_E(x_m, y_{m+1})\frac{\mathrm{d}\mu(x_m)}{\mathrm{d}x} \\ &+ \frac{\partial y_m}{\partial x_m}\int_{x_{m-1}}^{x_m} \frac{\partial \eta_E(x, y_m)}{\partial y_m}\mathrm{d}\mu \\ &+ \frac{\partial y_{m+1}}{\partial x_m}\int_{x_m}^{x_{m+1}} \frac{\partial \eta_E(x, y_{m+1})}{\partial y_{m+1}}\mathrm{d}\mu,\end{aligned}$$

since the reconstruction levels y_m, y_{m+1} are the only ones which depend on x_m,

$$y_m = \frac{\int_{x_{m-1}}^{x_m} x \mathrm{d}\mu}{\int_{x_{m-1}}^{x_m} \mathrm{d}\mu}, \quad y_{m+1} = \frac{\int_{x_m}^{x_{m+1}} x \mathrm{d}\mu}{\int_{x_m}^{x_{m+1}} \mathrm{d}\mu}.$$

The gradients $\frac{\partial y_m}{\partial x_m}, \frac{\partial y_{m+1}}{\partial x_m}$ are

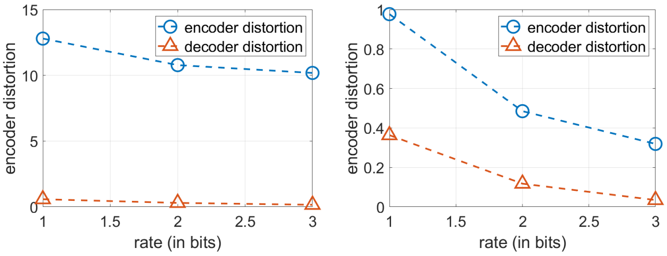
$$\frac{\partial y_m}{\partial x_m} = \frac{\mathrm{d}\mu(x_m)}{\mathrm{d}x}\frac{x_m - y_m}{\int_{x_{m-1}}^{x_m} \mathrm{d}\mu},$$

$$\frac{\partial y_{m+1}}{\partial x_m} = -\frac{\mathrm{d}\mu(x_m)}{\mathrm{d}x}\frac{x_m - y_{m+1}}{\int_{x_m}^{x_{m+1}} \mathrm{d}\mu}.$$

The codes are available at: https://tinyurl.com/GDnoiseless.

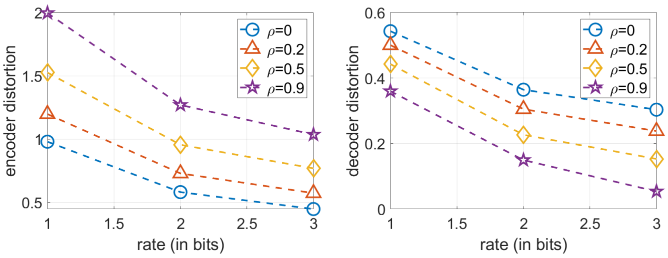
V. Numerical Results

We consider three different encoder distortion measures: i) $\eta_E = (x^3 - y)^2$, ii) $\eta_E = (3x/2 - y)^2$, and iii) $\eta_E = (x + \theta - y)^2$



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

Fig. 3: $X \sim \mathcal{N}(0,1), \eta_D = (x-y)^2$



(a) Encoder distortion (b) Decoder distortion

Fig. 4: $X \sim \theta \sim \mathcal{N}(0,1), \eta_E = (x+\theta-y)^2, \eta_D = (x-y)^2$

with $\eta_D = (x-y)^2$ for all three settings. We consider a Gaussian source $X \sim \mathcal{N}(0,1)$ for cases 1 and 2, and a jointly Gaussian source $(X,\theta) \sim \mathcal{N}\left(\begin{bmatrix}0\\0\end{bmatrix}, \begin{bmatrix}\sigma_X^2 & \rho\\ \rho & \sigma_\theta^2\end{bmatrix}\right), 0 \leq \rho < 1$ for case 3. Specifically, we present results for $\rho = [0, 0.2, 0.5, 0.9]$. We discretize θ for the third setting and compute a set of quantizers each corresponding to a realization of θ. We plot the encoder and decoder distortions in Figures 3a, 3b, and 4. We observe that the encoder's distortion increases with correlation as expected due to the strategic nature of the problem.

VI. Conclusion

In this paper, we have first formulated the strategic quantization problem and shown its features that differentiate it from classical quantization. We have numerically demonstrated that a strategic variation of Lloyd-Max may not converge to a local optimum. Motivated by this observation, we have developed gradient descent based solutions for the strategic quantization problem. Numerical results obtained via the proposed algorithm suggest several open theoretical questions about the behavior of the operational distortion-rate curve of the optimal strategic quantizers.

VII. Acknowledgement

This research is supported by the NSF via grants CCF #1910715 and CAREER #2048042.

Appendix

The decoder determines $\mathbf{y} = [y_1, \ldots, y_m]$ as follows:

$$y_m = \frac{\int_{x_{m-1}}^{x_m} x \mathrm{d}\mu}{\int_{x_{m-1}}^{x_m} \mathrm{d}\mu}.$$

The encoder's distortion and its derivative with respect to x_m:

$$J(\mathbf{q}) = \sum_{m=1}^{M} \int_{x_{m-1}}^{x_m} (x + \alpha - \beta y_m)^2 \mathrm{d}\mu,$$

$$\begin{aligned}\frac{\partial J}{\partial x_m} &= \frac{\mathrm{d}\mu(x_m)}{\mathrm{d}x}((x_m + \alpha - \beta y_m)^2 - (x_m + \alpha - \beta y_{m+1})^2) \\ &\quad - 2\beta \frac{\mathrm{d}y_m}{\mathrm{d}x_m} \int\limits_{x_{m-1}}^{x_m} (x + \alpha - \beta y_m) \mathrm{d}\mu \\ &\quad - 2\beta \frac{\mathrm{d}y_{m+1}}{\mathrm{d}x_m} \int\limits_{x_m}^{x_{m+1}} (x + \alpha - \beta y_{m+1}) \mathrm{d}\mu.\end{aligned}$$

Enforcing the KKT conditions for optimality

$$\frac{\partial J}{\partial x_m} = 0, \quad m \in [1:M] \tag{3}$$

we obtain, after some straightforward algebra, that the solutions that satisfies (3) are $\beta = 0, 2$, or $x_m = (y_m + y_{m+1})/2$ (the other condition is $y_{m+1} = y_m$ which is not possible since the actions are considered unique - if not, the corresponding regions could be combined). This implies that the quantizer is the same as the non-strategic quantizer if $\beta \neq 0, 2$, if the encoder decides to send something.

The distortion for a non-informative quantizer:

$$D_{nr} = \int_a^b (x + \alpha - \beta y)^2 \mathrm{d}\mu$$

The encoder's distortion can be re-written in terms of the non-revealing distortion and some other terms as

$$J = D_{nr} + \beta(\beta - 2)T, \tag{4}$$

where $T = \left(\sum\limits_{m=1}^{M} \frac{(\int_{x_{m-1}}^{x_m} x \mathrm{d}\mu)^2}{\int_{x_{m-1}}^{x_m} \mathrm{d}\mu} - \frac{(\sum\limits_{m=1}^{M} \int_{x_{m-1}}^{x_m} x \mathrm{d}\mu)^2}{\sum\limits_{m=1}^{M} \int_{x_{m-1}}^{x_m} \mathrm{d}\mu}\right)$. In order for the quantizer to be informative ($M > 1$), the second term has to be less than 0. This happens in three cases:

1) $\beta < 0$ and $T < 0$
2) $0 < \beta < 2$ and $T > 0$
3) $\beta > 2$ and $T < 0$

From Cauchy-Schwarz inequality, we have that for real numbers $u_1, u_2, \ldots, u_n$ and positive real numbers $v_1, v_2, \ldots, v_n$:

$$(\sum_{i=1}^{n} u_i)^2 / (\sum_{i=1}^{n} v_i) \leq \sum_{i=1}^{n} \frac{u_i^2}{v_i}.$$

Let $u_i = \int_{x_{m-1}}^{x_m} x \mathrm{d}\mu, v_i = \int_{x_{m-1}}^{x_m} \mathrm{d}\mu$. We get $T > 0$. This implies that the only possibility is case 2 with $0 < \beta < 2$, and the encoder chooses a non-strategic quantizer (as we show earlier that the only solution when $\beta \neq 0, 2$ is a non-strategic encoder if the encoder sends some message). From (4), we see that for $\beta = 0, 2$ the encoder distortion is the same as non-revealing distortion regardless of the quantizer used.

The optimal policy for the encoder is to be fully revealing for $\beta \in (0, 2)$, and the distortion remains the same for any M level quantization when $\beta = 0, 2$, and non-revealing otherwise.

REFERENCES

[1] L. Rayo and I. Segal, “Optimal Information Disclosure,” *Journal of Political Economy*, vol. 118, no. 5, pp. 949–987, 2010.

[2] E. Kamenica and M. Gentzkow, “Bayesian Persuasion,” *American Economic Review*, vol. 101, no. 6, pp. 2590–2615, 2011.

[3] M. Ostrovsky and M. Schwarz, “Information Disclosure and Unraveling in Matching Markets,” *American Economic Journal: Microeconomics*, vol. 2, no. 2, pp. 34–63, 2010.

[4] R. Alonso and O. Câmara, “Persuading Voters,” *American Economic Review*, vol. 106, no. 11, pp. 3590–3605, 2016.

[5] O. Candogan and K. Drakopoulos, “Optimal Signaling of Content Accuracy: Engagement vs. Misinformation,” *Operations Research*, vol. 68, no. 2, pp. 497–515, 2020.

[6] E. Akyol, C. Langbort, and T. Başar, “Privacy Constrained Information Processing,” in *54th IEEE conference on decision and control (CDC)*. IEEE, 2015, pp. 4511–4516.

[7] ——, “Information-Theoretic Approach to Strategic Communication as a Hierarchical Game,” *Proceedings of the IEEE*, vol. 105, no. 2, pp. 205–218, 2016.

[8] E. Kamenica, “Bayesian Persuasion and Information Design,” *Annual Review of Economics*, vol. 11, pp. 249–272, 2019.

[9] D. Bergemann and S. Morris, “Information Design: A Unified Perspective,” *Journal of Economic Literature*, vol. 57, no. 1, pp. 44–95, 2019.

[10] S. Dughmi and H. Xu, “Algorithmic Bayesian Persuasion,” *SIAM Journal on Computing*, no. 0, pp. STOC16–68, 2019.

[11] Y. C. Aybaş and E. Türkel, “Persuasion with Coarse Communication,” *arXiv preprint arXiv:1910.13547*, 2019.

[12] R. M. Gray and D. L. Neuhoff, “Quantization,” *IEEE Transactions on Information Theory*, vol. 44, no. 6, pp. 2325–2383, 1998.

[13] V. P. Crawford and J. Sobel, “Strategic Information Transmission,” *Econometrica: Journal of the Econometric Society*, pp. 1431–1451, 1982.

[14] H. Kono and M. Kandori, “Corrigendum to Crawford and Sobel (1982)“Strategic Information Transmission”,” *Econometrica*, vol. 89, no. 4, pp. 1–10, 2021.

[15] J. Sobel, “Giving and Receiving Advice,” *Advances in Economics and Econometrics*, vol. 1, pp. 305–341, 2013.

[16] A. Gersho and R. M. Gray, *Vector Quantization and Signal Compression*. Springer Sci. & Business Media, 2012, vol. 159.

[17] P. Dworczak and G. Martini, “The Simple Economics of Optimal Persuasion,” *Journal of Political Economy*, vol. 127, no. 5, pp. 1993–2048, Oct. 2019.

[18] J. Mensch, “Monotone Persuasion,” *Games and Economic Behavior*, vol. 130, pp. 521–542, 2021. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0899825621001263