Strategic Quantization of a Noisy Source

Anju Anand and Emrah Akyol

***Abstract*— This paper is concerned with strategic quantization of a noisy source where the encoder, which observes the remote source through a noisy channel, and the decoder, with a distortion defined over the remote source, have misaligned objectives. Such scenarios constitute a special class of games called information design in Economics. Here, we approach them with an Engineering lens by adding a constraint on the message space in the problem setting. This is an extension of our recent work on strategic quantization of a noiseless source. Finally, we present the numerical results that validate the analysis in this paper. The codes are available at: `https://tinyurl.com/allerton2023`.**

I. INTRODUCTION

In this paper, we study the quantizer design problem for the setting where an encoder that observes the source through a noisy channel, and a decoder with misaligned objectives communicate over a noiseless channel.

This problem in its conventional setting of identical objectives dates back to the seminal work of Dobrushin and Lyba [1], has been well studied in the literature since, see e.g., [2]–[4]. The main result of these prior works is that one can transform the problem of indirect source coding to a direct source coding problem with a modified distortion measure defined as the original distortion conditioned over the sensing channel output.

Our problem here is closely related to a class of communication games known as "information design," also known as "Bayesian Persuasion," where agents with diverging objectives communicate as detailed in Section II.B.

The problem setting has several applications in engineering as well as Economics. For an engineering application, consider the Internet of Things, where agents with misaligned objectives communicate over channels with delay constraints. For a more concrete, real-life application, consider two smart cars by competing manufacturers, e.g., Tesla and Honda, where the Tesla (decoder) car asks for a piece of specific information, such as traffic congestion, from the Honda (encoder) to decide on changing its route or not. Say Honda's objective is to make Tesla take a specific action, e.g., to change its route, while Tesla's objective is to estimate the congestion to make the right decision accurately. Honda's objective is obviously different from that of Tesla, hence has no incentive to convey a truthful congestion estimate. However, Tesla is aware of Honda's motives while still would like to use Honda's information (if possible). With the realistic assumption that Honda would observe this information through a noisy sensing channel (i.e., a sensor), how would these cars communicate over a fixed-rate zero-delay channel? Such problems can be handled using our model. Note that here Honda has three different behavioral choices: it can choose not to communication (non-revealing strategy), can communicate exactly what the Tesla wants (fully-revealing strategy), or it can craft a message that would make Tesla to change its route. Note that Tesla can choose not to use Honda's message, if it is statistically too far from the truth. Hence, crafting an optimal message for Honda that would serve its own objective, knowing that Tesla's objective differs from it, is a formidable research challenge.

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

A. Anand is with the Department of Electrical and Computer Engineering, Binghamton University, SUNY, USA aanand6@binghamton.edu

E. Akyol is with Faculty of Electrical and Computer Engineering, Binghamton University, SUNY, USA eakyol@binghamton.edu

II. PRELIMINARIES

A. Notation

In this paper, random variables are denoted by capital letters, their sample values are denoted by the respective lower case letters, and their alphabets are denoted by the respective calligraphic letters. This alphabet may be finite, countably infinite, or a continuum, like an interval $[a, b] \subset \mathbb{R}$. The expectation operator is denoted by $\mathbb{E}\{\cdot\}$. The scalar Gaussian with mean m, variance σ^2 is denoted by $\mathbb{N}(m, \sigma^2)$. All logarithms are base 2.

B. Strategic Quantization Prior Work

The strategic quantization problem can be described as follows: the encoder observes a signal $X \in \mathcal{X}$, and sends a message $Z \in \mathcal{Z}$ to the decoder, upon receiving which the decoder takes the action $Y \in \mathcal{Y}$. The encoder designs the quantizer decision levels Q to minimize its objective D_E, while the decoder designs the quantizer representative levels $\mathbf{y}$ to minimize its objective D_D. Note that the objectives of the encoder and the decoder are misaligned ($D_E \neq D_D$). The strategic quantizer is a mapping $Q : \mathcal{X} \to \mathcal{Z}$, with $|\mathcal{Z}| \leq M$ for a given quantization resolution $M \in \mathbb{Z}^+$, and given distortion measures D_E, D_D.

As mentioned earlier, our problem is a variation of the Bayesian Persuasion (or information design) class of problems where an encoder and decoder with misaligned objectives communicate [5]. This class of the problems have been an active research area in Economics due to their modeling abilities of real-life scenarios, see e.g., [6]–[9].

This problem was previously studied in Economics as well as Computer Science. In [10], authors showed the existence of optimal strategic quantizers in abstract spaces. Moreover, authors provide a low-complexity method to obtain the optimal strategic quantizer. In [11], [12], authors characterize sufficient conditions for the monotonicity of the optimal strategic quantizer, and as a byproduct of their analysis, characterize its behavior (non-revealing, fully revealing, or partially revealing) for some special settings. In Computer Science, in [13], this problem was studied from a computational perspective and they report approximate results on this problem, relating to another problem they solved conclusively. One of their main results is that they showed the algorithmic complexity of finding the optimum strategic quantizer as NP-hard.

In [14], we showed that a strategic variation of the Lloyd-Max algorithm does not converge to a locally optimal solution. As a remedy we develop a gradient descent based solution for this problem. We also demonstrated that even for well-behaving sources, such as scalar Uniform, there are multiple local optima, depending on the distortion measures chosen, in sharp contrast with the classical quantization for which the local optima is unique for the case of log-concave sources (which includes Uniform sources). We also analyzed the behavior of the optimal strategic quantizer for some typical settings. The behavior can be one of the following three: i) Non-revealing: the encoder does not send any information, i.e., $Q(X) =$ constant. ii) Fully revealing: the encoder effectively sends the information the decoder asks, which simplifies the problem into classical quantizer design with the decoder's objective. iii) Partially revealing: the encoder sends some information but not exactly what the decoder wants.

In [15], [16], we carried out our analysis of strategic quantization to the scenario where there is a noisy communication channel between the encoder and the decoder, using random index mapping in conjunction with gradient descent based and dynamics programming solutions respectively. In [17], we derived the globally optimal strategic quantizer via a dynamic programming based solution to resolve the poor local minima issues with gradient descent based solutions.

In Appendix I, we prove the following result, which is an extension of a result presented in [14], as well as in [11]:

Theorem 1. *For $\eta_E(u, y) = (\alpha u - y)^2$ and $\eta_D(u, y) = (u - y)^2$, the optimal strategic quantizer Q is:*

$$Q(x) = \begin{cases} \arg\min \mathbb{E}\{(\hat{U} - Q(X))^2\}, & \text{for } 0 < \alpha < 1/2 \\ arbitrary, & \text{for } \alpha \in \{0, 1/2\} \\ constant, & \text{otherwise} \end{cases}$$

where $\hat{U} = \mathbb{E}\{U|X\}$. Note that the first case corresponds to the fully-revealing behavior, while the second one is non-revealing.

We refer to this theorem, later in the text, in order to demonstrate the use of our main result in this paper.

C. Remote Source Coding Prior Work

As mentioned earlier, this problem is well-studied in the classical, i.e., non-strategic, compression literature, under different names such as remote source coding, indirect rate-distortion, noisy quantization etc.

The main result, by Dobrushin and Tsybakov [1], adopted to the quantization setting as in [3] is presented as follows:

Theorem 2. *Consider the remote source coding problem where the source U is observed through a memoryless channel $P(X|U)$ by the encoder. The encoder quantizes the channel output, X, to minimize a common distortion measure $\mathbb{E}\{d(U, Q(X))\}$ subject to a rate constraint. Let Q_1 be the optimal quantizer, i.e.,*

$$Q_1 = \arg\min \mathbb{E}\{d(U, Q(X))\}$$

Let Q_2 be the optimal point-to-point quantizer for the distortion metric $d_2(C, b) = \mathbb{E}\{d(A, b)|C\}$ where $P_{U|X} = P_{A|C}$, i.e.,

$$Q_2 = \arg\min \mathbb{E}\{d_2(U, Q(U)\}$$

where d_2 is defined as above.

Then, $Q_1 = Q_2$.

D. Problem Definition

Consider the following quantization problem: an encoder observes a realization of a scalar source $U \in \mathcal{U}$ as $X \in \mathcal{X}, X = U + W$, where W is an additive noise independent of the source with a probability distribution μ_W. The joint probability distribution of the source and its noisy version is given by $\mu_{U,X}$. The encoder maps X 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, $Q : \mathcal{X} \to \mathcal{Z}$. After receiving the message Z, the decoder applies a mapping $\phi : \mathcal{Z} \to \mathcal{Y}$, where $|\mathcal{Y}| = |\mathcal{Z}|$, on the message Z and takes an action $Y = \phi(Z)$. The encoder and decoder minimize their respective objectives $D_E = \mathbb{E}\{\eta_E(U, Y)\}$ and $D_D = \mathbb{E}\{\eta_D(U, Y)\}$, 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, and the statistics of the source $\mu_{U,X}(\cdot, \cdot)$. The objectives (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 for the equilibrium, i.e., the encoder minimizes its distortion if used with a corresponding decoder that minimizes its own distortion. This communication setting is given in Figure 1.

III. MAIN RESULTS

A. Analysis

The objectives of the encoder and the decoder are given by $D_s = \mathbb{E}\{\eta_s(U, Q(X))\}, s \in \{E, D\}$. The set $\mathcal{X}$ is divided into mutually exclusive and exhaustive sets $\mathcal{V}_1, \mathcal{V}_2, \ldots, \mathcal{V}_M$. We make the following "monotonicity" assumption.

Assumption 3 (Convex code-cells)**.** *$\mathcal{V}_i$ is convex for all $i \in [1 : M]$.*

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

$$\mathcal{V}_i = [x_{i-1}, x_i).$$

Problem 1. *The source $U \in \mathcal{U}$ is corrupted by an independent additive noise $W \sim \mu_W$, and is observed by the encoder as $X = U + W$ with a joint probability distribution $\mu_{U,X}$. The encoder communicates to the decoder over a noiseless channel with rate R. The objectives of the encoder are misaligned and are given by η_E and η_D, respectively, with $\eta_E \neq \eta_D$. Find the quantizer decision levels Q, and the set of actions $\mathbf{y} = [y_1, \dots, y_M]$ as a function of the quantizer decision levels that satisfy:*

$$q^* = \arg\min_{q} \sum_{i=1}^{M} \mathbb{E}\{\mathbb{E}\{\eta_E(u, \mathbf{y}) | x \in \mathcal{V}_i\},$$

where actions $\mathbf{y}$ are $y_i^* = \arg\min_{y_i \in \mathcal{Y}} \mathbb{E}\{\eta_D(u, \mathbf{y}) | x \in \mathcal{V}_i\} \forall i \in [1:M]$, and the rate satisfies $\log M \leq R$.

All integrals are over $\mathcal{V}_i$, unless specified otherwise. The encoder chooses the quantizer Q with boundary levels $[x_0, \dots, x_M]$. The decoder determines a set of actions $\mathbf{y} = [y_1, \dots, y_M]$ as the best response to Q to minimize its cost D_D as

$$y_i^* = \arg\min_{y_i \in \mathcal{Y}} \sum_{i=1}^{M} \mathbb{E}\{\eta_D(u, \mathbf{y}) | x \in \mathcal{V}_i\}.$$

After observing x, the encoder quantizes the source as

$$z_i = Q(x), \quad x \in \mathcal{V}_i.$$

The decoder receives the message z_i transmitted over a noiseless channel and takes the action

$$y_i = \phi(z_i).$$

The distortions to the encoder and the decoder and the optimum decoder reconstruction for $i \in [1:M]$ are

$$D_s = \sum_{i=1}^{M} \int \int_{u \in \mathcal{U}} \eta_s(u, y_i) \mathrm{d}\mu_{U,X},$$

$$y_i = \arg\min_{y \in \mathcal{Y}} \int \int_{u \in \mathcal{U}} \eta_D(u, y) \mathrm{d}\mu_{U,X}.$$

The reconstruction levels $\mathbf{y}$ are found using KKT conditions,

$$\frac{\partial D_D}{\partial y_i} = \int \int_{u \in \mathcal{U}} \frac{\partial}{\partial y_i} \eta_D(u, y_i) \mathrm{d}\mu_{U,X}.$$

For $\eta_D(u, y) = (u - y)^2$, we have

$$\frac{\partial D_D}{\partial y_i} = -2 \int \int_{u \in \mathcal{U}} (u - y_i) \mathrm{d}\mu_{U,X},$$

$$y_i = \frac{\int \int_{u \in \mathcal{U}} u \mathrm{d}\mu_{U,X}}{\int \int_{u \in \mathcal{U}} \mathrm{d}\mu_{U,X}}.$$

B. Main Result

Theorem 4. *The noisy strategic quantization problem described above, with distortions $\eta_E(u, y)$ and $\eta_D(u, y)$ is equivalent to the noiseless strategic quantization problem with a modified encoder distortion measure $\eta'_E(x, y) = \mathbb{E}\{\eta_E(u, y) | X = x\}$ for a given $P(X|U)$ observation channel.*

Proof:
The encoder distortion D_E can be written in terms of only the noisy source realization available to the encoder as

$$\begin{aligned} D_E &= \mathbb{E}\{\eta_E(U, Q(X))\} = \mathbb{E}\{\mathbb{E}\{\eta_E(U, Q(X)) | X\}\} \\ &= \mathbb{E}\{\eta'_E(X, Q(X))\} \end{aligned}$$

where $\eta'_E(X, Q(X)) = \mathbb{E}\{\eta_E(U, Q(X)) | X\}$ and decoder reconstruction

$$y_i^* = \arg\min_{y_i \in \mathcal{Y}} \mathbb{E}\{\eta_D(u, y_i) | x \in \mathcal{V}_i\}$$

C. Quadratic Gaussian Setting

Let $\eta_E(u, \theta, y) = (u + \theta - y)^2, \eta_D(u, \theta, y) = (u - y)^2$. We extend the previous quantizer design to 2−dimensional quantization by designing a set of quantizers each corresponding to a realization of θ, $Q = \{Q_\theta | \theta \in \mathcal{T}\}$, where Q_θ is a quantizer with decision levels $[v_{\theta,0}, \dots, v_{\theta,M}]$. Similar to the 1-dimensional setting, we make a monotonicity assumption here as follows.

Assumption 5 (Convex code-cells)**.** *$\mathcal{V}_{\theta,i}$ is convex for all $i \in [1:M]$.*

Under assumption 1, $\mathcal{V}_{\theta,i}$ is an interval since X is a scalar, i.e., $\mathcal{V}_{\theta,i} = [x_{\theta,i-1}, x_{\theta,i})$. All integrals in this subsection are over $\mathcal{V}_{\theta,i}$, unless specified otherwise. The encoder designs Q to minimize

$$D_E = \sum_{i=1}^{M} \int_{\theta \in \mathcal{T}} \int \int_{u \in \mathcal{U}} \eta_E(u, \theta, y_i) \mathrm{d}\mu_{U,X,\theta},$$

where the optimal representative levels y_i are computed by the decoder by minimizing its distortion $\mathbb{E}\{\eta_D(U, \theta, Q(X))\}$,

$$y_i = \arg\min_{y \in \mathcal{Y}} \int_{\theta \in \mathcal{T}} \int \int_{u \in \mathcal{U}} \eta_D(u, \theta, y) \mathrm{d}\mu_{U,X,\theta}.$$

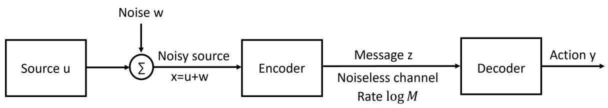


Fig. 1: Communication diagram

When $\eta_D = (u-y)^2$,

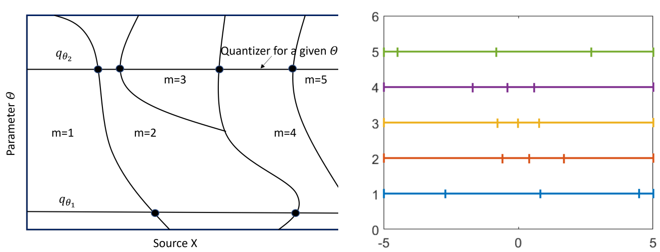
$$\mathbb{E}\{\eta_D(U,\theta,Q(X))\} = \sum_{i=1}^{M} \int_{\theta\in\mathcal{T}} \int \int_{u\in\mathcal{U}} (u-y_i)^2 \mathrm{d}\mu_{U,X,\theta},$$

$$y_i = \arg\min_{y\in\mathcal{Y}} \int_{\theta\in\mathcal{T}} \int \int_{u\in\mathcal{U}} (u-y)^2 \mathrm{d}\mu_{U,X,\theta}.$$

KKT optimality conditions imply

$$y_i = \frac{\int_{\theta\in\mathcal{T}} \int \int_{u\in\mathcal{U}} u\mathrm{d}\mu_{U,X,\theta}}{\int_{\theta\in\mathcal{T}} \int \int_{u\in\mathcal{U}} \mathrm{d}\mu_{U,\theta,X}} = \mathbb{E}\{U|X\in\mathcal{V}_{:,i}\}.$$

We show in Figure 2a that the nature of the quantizer may change with the value of θ. For the two quantizers shown, we see that the rate is $\log 5$ and $\log 3$. In Figure 2b, we show an example quantizer for θ discretized to 5 values.



(a) M=5 level quantizer (b) Example quantizer

Fig. 2: Quantization of X parameterized by θ.

We re-write the objective function D_E in terms of distortion due to the noisy source and distortion due to quantization. The term $Q(X)$ is written as Q for brevity. Note that $U-X-Q(X)-Y$, and $\theta-X-Q(X)$ forms a markov chain, with $Y=\mathbb{E}\{U|Q\}$.

$$\begin{aligned}
D_E &= \mathbb{E}\{(U+\theta-Y)^2\}\\
&= \mathbb{E}_X\{\mathbb{E}\{(U+\theta-Y)^2|X\}\}\\
&= \mathbb{E}_X\{\mathbb{E}\{(U-\mathbb{E}\{U|X\}+\mathbb{E}\{U|X\}-\theta-Y)^2|X\}\}\\
&= \mathbb{E}_X\{\mathbb{E}\{(U-\mathbb{E}\{U|X\})^2+(\mathbb{E}\{U|X\}-\theta-Y)^2\\
&\quad + 2(U-\mathbb{E}\{U|X\})(\mathbb{E}\{U|X\}-\theta-Y)|X\}\}\\
&= \mathbb{E}\{(U-\mathbb{E}\{U|X\})^2\}+\mathbb{E}\{(\mathbb{E}\{U|X\}-\theta-Y)^2\}\\
&\quad + 2\mathbb{E}_X\{\mathbb{E}\{(U-\mathbb{E}\{U|X\})(\mathbb{E}\{U|X\}-\theta-Y)|X\}\}
\end{aligned}$$

The terms $\mathbb{E}\{(U-\mathbb{E}\{U|X\})\mathbb{E}\{U|X\}|X\}$ and $\mathbb{E}\{(U-\mathbb{E}\{U|X\})Y|X\}$ both vanish due to the orthogonality principle in optimal estimation. The third term,

$$\mathbb{E}\{(U-\mathbb{E}\{U|X\})\theta|X\} = \mathbb{E}\{U\theta|X\} - \mathbb{E}\{U|X\}\mathbb{E}\{\theta|X\}.$$

Assuming U and θ independent, we have

$$D_E = \mathbb{E}\{(U-\mathbb{E}\{U|X\})^2\}+\mathbb{E}\{(\mathbb{E}\{U|X\}-\theta-Y)^2\}.$$

Minimizing $D_E = \mathbb{E}\{d_1(X,\theta,Q(X))\}$ is equivalent to minimizing $D'_E = \mathbb{E}\{d_1^*(X,\theta,Y)\}$, where

$$d_1^*(X,\theta,Y) = \mathbb{E}\{(\mathbb{E}\{U|X\}+\theta-Y)^2\},$$

since the other term $\mathbb{E}\{(U-\mathbb{E}\{U|X\})^2\}$ is distortion due to the noise at the source, which cannot be optimized. The encoder minimizes its equivalent distortion

$$\begin{aligned}
D'_E &= \mathbb{E}\Big\{\sum_{i=1}^{M}(\mathbb{E}\{U|X\}+\theta-y_i)^2|\mathcal{V}_i\Big\}\\
&= \sum_{i=1}^{M}\int_{u\in\mathcal{U}}\int_{\theta\in\mathcal{T}}\int(\mathbb{E}\{U|X\}+\theta-y_i)^2\mathrm{d}\mu_{U,X,\theta},
\end{aligned}$$

where

$$y_i = \frac{\int_{\theta\in\mathcal{T}} \int \int_{u\in\mathcal{U}} u\mathrm{d}\mu_{U,X,\theta}}{\int_{\theta\in\mathcal{T}} \int \int_{u\in\mathcal{U}} \mathrm{d}\mu_{U,X,\theta}} = \mathbb{E}\{U|X\in\mathcal{V}_{:,i}\}.$$

D. Algorithm

The problem setting requires the encoder to choose decision levels first, followed by the decoder's choice of reconstruction points. This allows a gradient descent based solution where the optimization parameter is the encoder's decision levels Q.

Starting with an arbitrary initialization of the quantizer boundary levels $Q=Q_0$, the reconstruction levels $\mathbf{y}(Q)$ and the corresponding encoder distortion D_E are computed. Then, the following steps are iterated until convergence:

1) Compute the gradients $\left\{\frac{\partial D_E}{\partial x_m}\right\}$
2) Update the decision levels Q as $x_m \triangleq x_m - \Delta\frac{\partial D_E}{\partial x_m}$ if $\{x_m\}$ adheres to quantizer constraints.

We present below the derivation of the gradients for an MSE decoder with a quadratic encoder distortion $\eta_E = (u+\theta-y)^2$. The gradients of the encoder's distortion with respect to the decision levels are

$$\begin{aligned}
\frac{\partial D_E}{\partial x_{\theta',i}} &= \int_{u\in\mathcal{U}} (u+\theta'-y_i)^2\frac{\mathrm{d}\mu_{U,X,\theta}}{\mathrm{d}x\mathrm{d}\theta}(x_{\theta',i},\theta')\\
&\quad - \int_{u\in\mathcal{U}} (u+\theta'-y_{i+1})^2\frac{\mathrm{d}\mu_{U,X,\theta}}{\mathrm{d}x\mathrm{d}\theta}(x_{\theta',i},\theta')\\
&\quad - 2\int_{u\in\mathcal{U}}\int_{\theta\in\mathcal{T}}\int_{x_{\theta,i-1}}^{x_{\theta,i}}(u+\theta-y_i)\frac{\mathrm{d}y_i}{\mathrm{d}x_{\theta',i}}\mathrm{d}\mu_{U,X,\theta}\\
&\quad - 2\int_{u\in\mathcal{U}}\int_{\theta\in\mathcal{T}}\int_{x_{\theta,i}}^{x_{\theta,i+1}}(u+\theta-y_{i+1})\frac{\mathrm{d}y_{i+1}}{\mathrm{d}x_{\theta',i}}\mathrm{d}\mu_{U,X,\theta},
\end{aligned}$$

where

$$\frac{\mathrm{d}y_i}{\mathrm{d}x_{\theta',i}} = \frac{\int\limits_{u\in\mathcal{U}} u\frac{\mathrm{d}\mu_{U,X,\theta}(u,x_{\theta',i},\theta')}{\mathrm{d}x\mathrm{d}\theta} - y_i \int\limits_{u\in\mathcal{U}} \frac{\mathrm{d}\mu_{U,X,\theta}(u,x_{\theta',i},\theta')}{\mathrm{d}x\mathrm{d}\theta}}{\int\limits_{\theta\in\mathcal{T}} \int\limits_{x_{\theta,i-1}}^{x_{\theta,i}} \int\limits_{u\in\mathcal{U}} \mathrm{d}\mu_{U,X,\theta}},$$

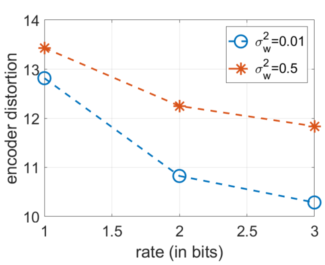
$$\frac{\mathrm{d}y_{i+1}}{\mathrm{d}x_{\theta',i}} = -\frac{\int\limits_{u\in\mathcal{U}} u\frac{\mathrm{d}\mu_{U,X,\theta}(u,x_{\theta',i},\theta')}{\mathrm{d}x\mathrm{d}\theta} - y_i \int\limits_{u\in\mathcal{U}} \frac{\mathrm{d}\mu_{U,X,\theta}(u,x_{\theta',i},\theta')}{\mathrm{d}x\mathrm{d}\theta}}{\int\limits_{\theta\in\mathcal{T}} \int\limits_{x_{\theta,i}}^{x_{\theta,i+1}} \int\limits_{u\in\mathcal{U}} \mathrm{d}\mu_{U,X,\theta}}.$$

IV. NUMERICAL RESULTS

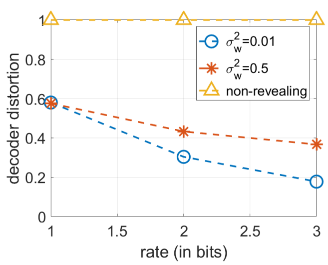
In this section, we present numerical results for two settings:

1) Source $X \sim \mathbb{N}(0,1), \eta_E(u,y) = (u^3-y)^2, \eta_D(u,y) = (u-y)^2$.
2) Quadratic-Gaussian setting with a 2-dimensional source $(U,\theta), U \sim \mathcal{N}(0,1)$, discretized $\theta \sim \mathcal{N}(0,1)$, U and θ independent, $\eta_E(u,\theta,y) = (u+\theta-y)^2, \eta_D(u,\theta,y) = (u-y)^2$.

in Figures 3 and 4 respectively, with an independent additive noise $W \sim \mathbb{N}(0,0.01)$ and $W \sim \mathbb{N}(0,0.5)$. We observe from Figures 3b and 4b that the decoder distortion for a non-revealing $M=1$ quantizer (or if the decoder chooses not to accept encoder's message) is greater than when the decoder acts according to the information from the encoder.

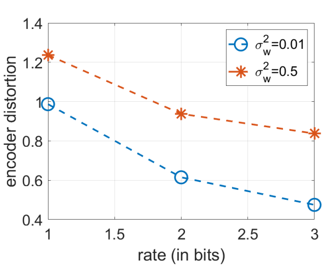


(a) Encoder distortion

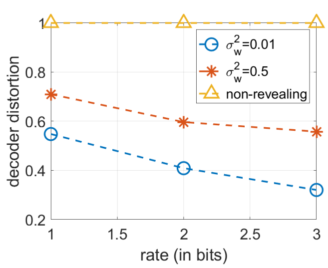


(b) Decoder distortion

Fig. 3: Source $U \sim \mathbb{N}(0,1)$ with $\eta_E(u,y) = (u^3-y)^2, \eta_D(u,y) = (u-y)^2$.



(a) Encoder distortion



(b) Decoder distortion

Fig. 4: Source $(U,\theta) \sim \mathbb{N}\left(\begin{bmatrix}0\\0\end{bmatrix}, \begin{bmatrix}1 & 0\\0 & 1\end{bmatrix}\right)$ with $\eta_E(u,\theta,y) = (u+\theta-y)^2, \eta_D(u,\theta,y) = (u-y)^2$.

V. CONCLUSIONS

In this paper, we extended our gradient descent based strategic quantizer of a noiseless source design approach to that for a noisy source. We considered a 1-dimensional and a 2-dimensional source. Obtained numerical results suggest the validity of the algorithm.

APPENDIX I
QUANTIZER BEHAVIOUR FOR $\eta_E = (u+\alpha-\beta y)^2, \eta_D = (u-y)^2$

Consider a continuous noisy source $X = U + N$, where N is the additive noise to the noiseless source U, $(U,X) \sim \mu_{U,X}$, $\eta_E(u,y) = (u+\alpha-\beta y)^2$, $\eta_D(u,y) = (u-y)^2$ for a given $\alpha, \beta \in \mathbb{R}$ quantized to M levels. In other words, the decoder wants to reconstruct U as closely as possible, while the encoder wants the decoder's construction to be as close as possible to $\frac{U+\alpha}{\beta}$, both in the MSE sense. Can the encoder "persuade" the decoder by carefully designing quantizer intervals $\mathcal{V}_m^*$?

The objective function is

$$J = \int\limits_{\mathcal{U}} \sum_{m=1}^{M} \int\limits_{x_{m-1}}^{x_m} (u+\alpha-\beta y_m)^2 \mathrm{d}\mu_{U,X},$$

where y_m is given by

$$y_m = \frac{\int\limits_{\mathcal{U}} \int\limits_{x_{m-1}}^{x_m} u\mathrm{d}\mu_{U,X}}{\int\limits_{\mathcal{U}} \int\limits_{x_{m-1}}^{x_m} \mathrm{d}\mu_{U,X}}$$

The derivative of the objective function with respect to quantizer decision level x_m,

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

where $\frac{\mathrm{d}y_m}{\mathrm{d}x_m}, \frac{\mathrm{d}y_{m+1}}{\mathrm{d}x_m}$ are

$$\begin{aligned}\frac{\mathrm{d}y_m}{\mathrm{d}x_m} &= \Bigg(\int\limits_{\mathcal{U}} u\frac{\mathrm{d}\mu_{U,X}(u,x_m)}{\mathrm{d}x} \int\limits_{\mathcal{U}} \int\limits_{x_{m-1}}^{x_m} \mathrm{d}\mu_{U,X} \\ &\quad - \int\limits_{\mathcal{U}} \frac{\mathrm{d}\mu_{U,X}(u,x_m)}{\mathrm{d}x} \int\limits_{\mathcal{U}} \int\limits_{x_{m-1}}^{x_m} u\mathrm{d}\mu_{U,X} \Bigg) \\ &\quad \Bigg/ \Bigg(\int\limits_{\mathcal{U}} \int\limits_{x_{m-1}}^{x_m} \mathrm{d}\mu_{U,X} \Bigg)^2\end{aligned}$$

$$= \frac{\int\limits_{\mathcal{U}} u \frac{\mathrm{d}\mu_{U,X}(u,x_m)}{\mathrm{d}x} - y_m \int\limits_{\mathcal{U}} \frac{\mathrm{d}\mu_{U,X}(u,x_m)}{\mathrm{d}x}}{\int\limits_{\mathcal{U}} \int\limits_{x_{m-1}}^{x_m} \mathrm{d}\mu_{U,X}}$$

$$\frac{\mathrm{d}y_{m+1}}{\mathrm{d}x_m} = -\Bigg(\int\limits_{\mathcal{U}} u \frac{\mathrm{d}\mu_{U,X}(u,x_m)}{\mathrm{d}x} \int\limits_{\mathcal{U}} \int\limits_{x_m}^{x_{m+1}} \mathrm{d}\mu_{U,X} - \int\limits_{\mathcal{U}} \frac{\mathrm{d}\mu_{U,X}(u,x_m)}{\mathrm{d}x} \int\limits_{\mathcal{U}} \int\limits_{x_m}^{x_{m+1}} u\mathrm{d}\mu_{U,X} \Bigg) \Bigg/ \Bigg(\int\limits_{\mathcal{U}} \int\limits_{x_m}^{x_{m+1}} \mathrm{d}\mu_{U,X} \Bigg)^2$$

$$= -\frac{\int\limits_{\mathcal{U}} u \frac{\mathrm{d}\mu_{U,X}(u,x_m)}{\mathrm{d}x} - y_{m+1} \int\limits_{\mathcal{U}} \frac{\mathrm{d}\mu_{U,X}(u,x_m)}{\mathrm{d}x}}{\int\limits_{\mathcal{U}} \int\limits_{x_m}^{x_{m+1}} \mathrm{d}\mu_{U,X}}$$

Enforcing the KKT conditions of optimality,

$$\frac{\partial J}{\partial x_m} = 0, \tag{1}$$

we obtain after some straightforward algebra, that the solution that satisfies 1 are $\beta = 0, 2$, or $\mathbb{E}\{U|X = x_m\} = \frac{y_m + y_{m+1}}{2}$ (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 $\beta \notin \{0, 2\}$, if the encoder decides to send something.

The encoder's distortion can be simplified as

$$J = \int\limits_{\mathcal{U}} \sum_{m=1}^{M} \int\limits_{x_{m-1}}^{x_m} (u + \alpha - \beta y_m)^2 \mathrm{d}\mu_{U,X}$$
$$= \int\limits_{\mathcal{U}} \int\limits_{a_X}^{b_X} (u)^2 \mathrm{d}\mu_{U,X} + \alpha^2 + 2\alpha(1-\beta) \int\limits_{\mathcal{U}} \int\limits_{a_X}^{b_X} u\mathrm{d}\mu_{U,X} + \beta(\beta - 2) \sum_{m=1}^{M} y_m \int\limits_{\mathcal{U}} \int\limits_{x_{m-1}}^{x_m} u\mathrm{d}\mu_{U,X}$$

The distortion for a non-informative quantizer D_n:

$$D_n = \int\limits_{\mathcal{U}} \int\limits_{a_X}^{b_X} (u + \alpha - \beta y)^2 \mathrm{d}\mu_{U,X}$$
$$= \int\limits_{\mathcal{U}} \int\limits_{a_X}^{b_X} u^2 \mathrm{d}\mu_{U,X} + \alpha^2 + 2\alpha(1-\beta) \int\limits_{\mathcal{U}} \int\limits_{a_X}^{b_X} u\mathrm{d}\mu_{U,X} + \beta(\beta-2) y \int\limits_{\mathcal{U}} \int\limits_{a_X}^{b_X} u\mathrm{d}\mu_{U,X}.$$

The encoder distortion can be written in terms of D_n as

$$J = D_n + \beta(\beta-2) \Bigg(\sum_{m=1}^{M} y_m \int\limits_{\mathcal{U}} \int\limits_{x_{m-1}}^{x_m} u\mathrm{d}\mu_{U,X} - y \int\limits_{\mathcal{U}} \int\limits_{a_X}^{b_X} u\mathrm{d}\mu_{U,X} \Bigg)$$
$$= D_n + \beta(\beta-2) \Bigg(\sum_{m=1}^{M} \frac{\int\limits_{\mathcal{U}} \int\limits_{x_{m-1}}^{x_m} u\mathrm{d}\mu_{U,X}}{\int\limits_{\mathcal{U}} \int\limits_{x_{m-1}}^{x_m} \mathrm{d}\mu_{U,X}} \int\limits_{\mathcal{U}} \int\limits_{x_{m-1}}^{x_m} u\mathrm{d}\mu_{U,X} - \frac{\sum\limits_{m=1}^{M} \int\limits_{\mathcal{U}} \int\limits_{x_{m-1}}^{x_m} u\mathrm{d}\mu_{U,X}}{\sum\limits_{m=1}^{M} \int\limits_{\mathcal{U}} \int\limits_{x_{m-1}}^{x_m} \mathrm{d}\mu_{U,X}} \sum_{m=1}^{M} \int\limits_{\mathcal{U}} \int\limits_{x_{m-1}}^{x_m} u\mathrm{d}\mu_{U,X} \Bigg)$$
$$= D_n + \beta(\beta-2)\xi$$

where $\xi = \sum\limits_{m=1}^{M} \frac{\left(\int\limits_{\mathcal{U}} \int\limits_{x_{m-1}}^{x_m} u\mathrm{d}\mu_{U,X} \right)^2}{\int\limits_{\mathcal{U}} \int\limits_{x_{m-1}}^{x_m} \mathrm{d}\mu_{U,X}} - \frac{\left(\sum\limits_{m=1}^{M} \int\limits_{\mathcal{U}} \int\limits_{x_{m-1}}^{x_m} u\mathrm{d}\mu_{U,X} \right)^2}{\sum\limits_{m=1}^{M} \int\limits_{\mathcal{U}} \int\limits_{x_{m-1}}^{x_m} \mathrm{d}\mu_{U,X}}$.

The first term is the distortion for a non-informative quantizer ($M = 1$). In order for the quantizer to be informative, the second term has to be less than 0. This happens in three cases

1) $\beta < 0$ and $\xi < 0$,
2) $0 < \beta < 2$ and $\xi > 0$,
3) $\beta > 2$ and $\xi < 0$.

From Cauchy-Shwarz inequality $| < \mathbf{u}, \mathbf{v} > | \leq ||\mathbf{u}||||\mathbf{v}||$, substituting $u_i' = u_i / \sqrt{v_i}$ and $v_i' = \sqrt{v_i}$), we have that for real numbers $u_1, u_2, \ldots, u_n$ and positive real numbers $v_1, v_2, \ldots, v_n$:

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

In our case, $u_i = \int\limits_{\mathcal{U}} \int\limits_{x_{m-1}}^{x_m} u\mathrm{d}\mu_{U,X}, v_i = \int\limits_{\mathcal{U}} \int\limits_{x_{m-1}}^{x_m} \mathrm{d}\mu_{U,X}$ with real u_i and positive real v_i.

Applying this result to our problem, we obtain $\xi \leq 0$ which implies that the only possible case is case 2 with $0 < \beta < 2$. For $\beta \in \{0, 2\}$, the quantizer is arbitrary since encoder distortion is always D_n, and the encoder is a non-strategic quantizer and non-revealing for $\beta \notin [0, 2]$.

REFERENCES

[1] R. Dobrushin and B. Tsybakov, "Information Transmission with Additional Noise," *IRE Transactions on Information Theory*, vol. 8, no. 5, pp. 293–304, 1962.

[2] H. Witsenhausen, "Indirect Rate Distortion Problems," *IEEE Transactions on Information Theory*, vol. 26, no. 5, pp. 518–521, 1980.

[3] E. Ayanoglu, "On optimal quantization of noisy sources," *IEEE Transactions on Information Theory*, vol. 36, no. 6, pp. 1450–1452, 1990.

[4] J. Wolf and J. Ziv, "Transmission of Noisy Information to a Noisy receiver with Minimum Distortion," *IEEE Transactions on Information Theory*, vol. 16, no. 4, pp. 406–411, 1970.

[5] E. Kamenica and M. Gentzkow, "Bayesian Persuasion," *American Economic Review*, vol. 101, no. 6, pp. 2590–2615, 2011.
[6] D. Bergemann and S. Morris, "Information Design: A Unified Perspective," *Journal of Economic Literature*, vol. 57, no. 1, pp. 44–95, 2019.
[7] J. Sobel, "Giving and Receiving Advice," *Advances in Economics and Econometrics*, vol. 1, pp. 305–341, 2013.
[8] M. Ostrovsky and M. Schwarz, "Information Disclosure and Unraveling in Matching Markets," *American Economic Journal: Microeconomics*, vol. 2, no. 2, pp. 34–63, 2010.
[9] L. Rayo and I. Segal, "Optimal Information Disclosure," *Journal of Political Economy*, vol. 118, no. 5, pp. 949–987, 2010.
[10] Y. C. Aybaş and E. Türkel, "Persuasion with Coarse Communication," *arXiv preprint arXiv:1910.13547*, 2019.
[11] J. Mensch, "Monotone Persuasion," *Games and Economic Behavior*, vol. 130, pp. 521–542, 2021. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0899825621001263
[12] P. Dworczak and G. Martini, "The Simple Economics of Optimal Persuasion," *Journal of Political Economy*, vol. 127, no. 5, pp. 1993–2048, Oct. 2019.
[13] S. Dughmi and H. Xu, "Algorithmic Bayesian Persuasion," *SIAM Journal on Computing*, no. 0, pp. STOC16–68, 2019.
[14] E. Akyol and A. Anand, "Strategic Quantization," *IEEE International Symposium on Information Theory*, 2023. Available at https://tinyurl.com/GDnoiseless.
[15] A. Anand and E. Akyol, "Strategic Quantization over a Noisy Channel," 2023. Available at https://tinyurl.com/allerton2023.
[16] ——, "Channel-Optimized Strategic Quantizer Design via Dynamic Programming," *IEEE Statistical Signal Processing Workshop*, 2023. To appear, available at https://tinyurl.com/asilomar2023.
[17] ——, "Optimal Strategic Quantizer Design via Dynamic Programming," in *Proceedings of the IEEE Data Compression Conference*. IEEE, 2022, pp. 173–181.