Strategic Quantization with Quadratic Distortion Measures

Anju Anand, *Student Member, IEEE* and Emrah Akyol, *Senior Member, IEEE*

***Abstract*— We consider the problem of strategic quantization where an encoder and a decoder with misaligned objectives communicate over a rate-constrained noiseless channel. Specifically, we focus on a 2-dimensional source with quadratic distortion measures. We provide a design algorithm for this special case of strategic quantization, as well as an upper bound on the encoder distortion via employing linear communication strategies. Finally, we present comparative numerical results obtained via the proposed method, in conjunction with the aforementioned upper bound. We provide our numerical results and the code to obtain them for research purposes at `https://tinyurl.com/allerton24`.**

I. INTRODUCTION

Consider the following problem: Two smart cars by competing manufacturers, e.g., Tesla and Honda, are communicating, without sample delay, over a noiseless fixed bit rate channel. Tesla (the decoder) asks for traffic congestion information from Honda (the encoder), which is ahead in traffic, to decide on its route. Honda's objective might be to make Tesla take a specific action, e.g., change its current route, while Tesla wants to estimate the traffic conditions accurately. Since Honda's objective is different from Tesla's, Honda needs an incentive to convey a truthful congestion estimation. Tesla is aware of Honda's motives but would still like to use Honda's information. How would these cars communicate? Problems of this nature can be handled using the strategic quantization model (coarse persuasion) given in [1], [2], or more broadly, strategic communication models [3], [4]. Note that here, Honda has three different behavioral choices: it can choose not to communicate (non-revealing strategy), can precisely communicate what Tesla wants (fully-revealing strategy), or can craft a message that would make Tesla change its route (partially revealing strategy). 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 objective, knowing that Tesla's objective differs from it, is a formidable research challenge.

This research area has been well studied in Economics literature without the quantization cardinality constraint as the information design or the Bayesian persuasion problem [3], [5]. Such problems explore the use of information by a communication system designer (sender) to influence the action taken by a receiver [6], [7].

Strategic quantization was analyzed from a computational perspective in [2]. Aybaş and Türkel [1] studied the same problem via an information Economics lens, employing the mathematical tools developed in the Economics prior work, e.g., [3] and derived several theoretical properties of optimal strategic quantizers in general probability spaces. In [8], authors consider a Bayesian persuasion problem with an imperfect channel and a limited number of messages, and provide an upper bound on the pay-off to the sender. In [9], authors study Bayesian signalling games and characterize the minimum number of distinct source symbols that can be correctly recovered by a receiver in any equilibrium of this game, which they call the informativeness of the sender. In our prior work [10]–[13], we used the rich collection of prior quantizer design and optimization work to study this practically significant problem via an engineering lens. More specifically, in [11], we derived several properties of strategic quantization and proposed a straightforward gradient descent-based design strategy that yields a locally optimal strategic quantizer. In [10], we proposed a dynamic programming solution that achieves global optimality at the cost of increased complexity. In [12]–[14], we explored strategic quantization in noisy scenarios.

In a related but distinctly different class of signaling games called cheap talk [15], authors noted that quantizers can arise as equilibrium strategies endogenously, without an external constraint. In [15], 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. This results in a Nash equilibrium since both agents form a strategy that is the best response to each other's mapping because of the encoder's lack of commitment power in the cheap talk setting. In contrast, in the strategic quantization problem (and the information design problems in general as in [3], [5]), the encoder designs Q *ex-ante*, *before* seeing the source realization, and is committed to the designed Q afterward. This difference in commitment manifests in the notion of equilibrium since the encoder may not necessarily form a best response to the decoder due to its commitment to Q.

In this paper, we focus on the setting where both communicating agents use quadratic distortion measures. Particularly, the encoder observes a two-dimensional source $X, \theta \sim f(x, \theta)$ with a known joint distribution over X and θ, where X and θ can be interpreted as the state and bias variables. The decoder's objective is to estimate the state in the minimum mean squared error (MMSE) sense, i.e., the decoder minimizes $\mathbb{E}\{(X-\hat{X})^2\}$ by choosing an action $\hat{X}$ which is the optimal MMSE estimate of x given the quantization index from the encoder $y = Q(x, \theta)$, hence $\hat{X} = \mathbb{E}\{X|Y = y\}$. In sharp contrast with the conventional

Authors are with the Binghamton University–SUNY, Binghamton, NY, 13902 USA {aanand6,eakyol}@binghamton.edu. This research is supported by the NSF via grants CCF #1910715 and CAREER #2048042.

Problem. *For a given rate R, 2−dimensional source (X, θ) with a probability distribution function $f(x, \theta)$ find the decision boundaries $q_\theta = [x_{\theta,0}, x_{\theta,1}, \ldots, x_{\theta,M}], \forall\theta \in \mathcal{T}$ and actions $\mathbf{y}(Q) = [y_1, \ldots, y_M]$, where $Q = \{q_\theta, \theta \in \mathcal{T}\}$ as a function of boundaries that satisfy:*

$$Q^* = \arg\min_Q \sum_{m=1}^{M} \mathbb{E}\{(X + \theta - y_m^*(Q))^2 | x \in \mathcal{V}_{\theta,m}\},$$

where actions $\mathbf{y}(Q)$ are given as

$$y_m^*(Q) = \arg\min_{y\in\mathcal{Y}} \mathbb{E}\left\{(X - y)^2 | x \in \mathcal{V}_{:,m}\right\} \quad \forall m \in [1 : M],$$

and the rate satisfies $\log M \leq R$.

quantization problem where the encoder chooses Q that minimizes $\mathbb{E}\{(X - \hat{X})^2\}$, in this setting the encoder's choice of quantization mapping Q minimizes a biased estimate, i.e., $\mathbb{E}\{(X + \theta - \hat{X})^2\}$. The objectives and the source distribution are common knowledge, available for both agents. Similar signaling problems with quadratic measures have been analyzed in the Economics literature [15]–[17].

Our contributions in this paper are as follows:

1) gradient-descent based optimization method to obtain the locally optimal strategic quantizer, extending our algorithm for scalar sources to 2-dimensional sources,
2) upper bound on the encoder's distortion.

The basic design problem, as studied in [11], focuses on scalar settings, hence optimal strategic quantization of a two-dimensional source poses a challenge in developing an algorithmic solution similar to the one in [11]. We circumvent this issue by designing a separate quantizer for each realization of θ for the encoder[1]. Essentially, the encoder first determines which quantizer to use via observing the realization of θ, and then sends the quantization index where X lies in for that specific quantizer associated with θ realization. The decoder does not know the aforementioned θ realization (and hence the quantizer chosen by the encoder).

This paper is organized as follows: In Section II we present the problem formulation. In Section III, we present a gradient-descent based algorithm to compute the strategic quantizer, and an upper bound on the encoder distortion. We provide numerical results in Section IV, and conclude in Section V.

II. Preliminaries

A. Notation

In this paper, random variables are denoted using capital letters (say X), their sample values with respective lowercase letters (x), and their alphabet with respective calligraphic letters ($\mathcal{X}$). The set of real numbers is denoted by $\mathbb{R}$. The alphabet, $\mathcal{X}$, can be finite, infinite, or a continuum, like an interval $[a, b] \subset \mathbb{R}$. The uniform distribution over an interval $[t_1, t_2]$, and the 2-dimensional jointly Gaussian distribution with mean $\begin{bmatrix} t_1 & t_2 \end{bmatrix}'$ and respective variances σ_1^2, σ_2^2 with a correlation ρ are denoted by $U[t_1, t_2]$, and $\mathcal{N}\left(\begin{bmatrix} t_1 \\ t_2 \end{bmatrix}, \begin{bmatrix} \sigma_1^2 & \sigma_1\sigma_2\rho \\ \sigma_1\sigma_2\rho & \sigma_2^2 \end{bmatrix}\right), 0 \leq \rho < 1,\ t_1, t_2 \in \mathbb{R}$, respectively. The expectation operator is written as $\mathbb{E}\{\cdot\}$. The operator $|\cdot|$ denotes the absolute value if the argument is a scalar real number and the cardinality if the argument is a set.

B. Problem Formulation

Consider the following quantization problem: an encoder observes realizations of the two sources $X \in \mathcal{X} \subseteq [a_X, b_X], \theta \in \mathcal{T} \subseteq [a_\theta, b_\theta]$, $a_X, b_X, a_\theta, b_\theta \in \mathbb{R}$ with joint probability distribution $(X, \theta) \sim f(x, \theta)$, and maps (X, θ) 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 a non-injective mapping parameterized by θ, $q_\theta : \mathcal{X} \to \mathcal{Z}$. After receiving the message Z, the decoder applies a mapping $\phi : \mathcal{Z} \to \mathcal{Y}$ on the message Z and takes an action $Y = \phi(Z)$.[2]

The encoder and decoder minimize their respective objectives $D_E = \mathbb{E}\{\eta_E(X, \theta, Y)\} = \mathbb{E}\{(X + \theta - Y)^2\}$ and $D_D = \mathbb{E}\{\eta_E(X, Y)\} = \mathbb{E}\{(X - Y)^2\}$, which are misaligned ($\eta_E \neq \eta_D$). The encoder designs $Q = \{q_\theta, \theta \in \mathcal{T}\}$ *ex-ante*, i.e., without the knowledge of the realization of (X, θ), using only the objectives D_E and D_D, and the statistics of the source $f(\cdot, \cdot)$. The objectives (D_E and D_D), the shared prior (f), 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 Fig. 1, and the problem is summarized in the box. Since the encoder chooses the quantization decision levels Q first, followed by the decoder choosing the quantization representative levels ($\mathbf{y}$), we look for a Stackelberg equilibrium.

The set $\mathcal{X}$ is divided into mutually exclusive and exhaustive sets parameterized by the realization of θ as $\mathcal{V}_{\theta,1}, \mathcal{V}_{\theta,2}, \ldots, \mathcal{V}_{\theta,M}$. The $m-$th quantization region is denoted by $\mathcal{V}_{:,m} = \{\mathcal{V}_{\theta,m}, \forall\theta \in \mathcal{T}\}$. The encoder chooses the

[1] In cases where θ is not purely discrete, we discretize it over a inform grid

[2] An example scenario is as follows: a professor asked to write a reference letter for their student might write a biased letter (either positively or negatively) depending on their opinion of the student which may include aspects irrelevant to the job. Here, θ is the personal bias of the professor, and X is an objective measure of the relevant skills. The bias θ may be correlated to the student's performance X.

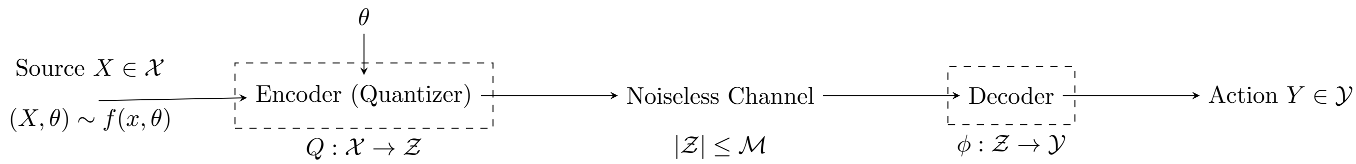

Fig. 1. Communication diagram: (X, θ) over a noiseless channel

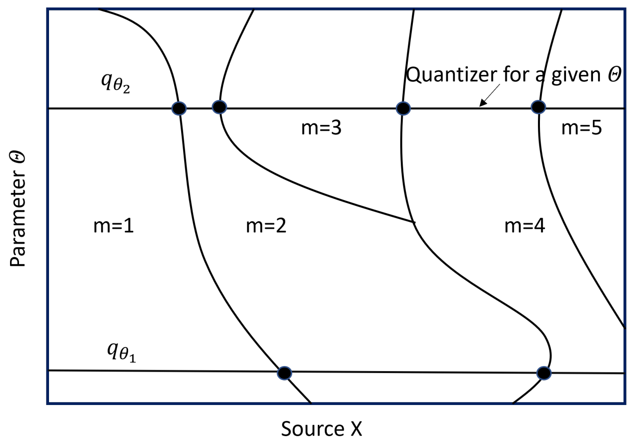

Fig. 2. Quantization of X parameterized by θ for $M = 5$ illustrated.

set of quantizers $Q = \{q_\theta, \theta \in \mathcal{T}\}$ to minimize its distortion,

$$D_E = \sum_{m=1}^{M} \int_{\theta \in \mathcal{T}} \int_{x \in \mathcal{V}_{\theta,m}} (x + \theta - y_m^*(Q))^2 \mathrm{d}f(x,\theta) \quad (1)$$

where the optimal reconstruction points y_m^* are determined by the decoder as a best response to Q to minimize its distortion,

$$\begin{aligned} y_m^* &= \arg\min_{y \in \mathcal{Y}} \sum_{m=1}^{M} \mathbb{E}\{(X-y)^2 | x \in \mathcal{V}_{:,m}\} \\ &= \frac{\int_{\theta \in \mathcal{T}} \int_{x \in \mathcal{V}_{\theta,m}} x \mathrm{d}f(x,\theta)}{\int_{\theta \in \mathcal{T}} \int_{x \in \mathcal{V}_{\theta,m}} \mathrm{d}f(x,\theta)}. \end{aligned} \quad (2)$$

The decoder determines a single set of actions $\mathbf{y}$ since it is unaware of the realization of θ.

Throughout this paper, we make the following "monotonicity" assumption on the sets $\{\mathcal{V}_{\theta,m}\}$.

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

Remark 1. *Assumption 1 is the first of the two regularity conditions commonly employed in the classical quantization literature, cf. [18]. Note that the second regularity condition, $y_m \in \mathcal{V}_m$, is not included in Assumption 1.*

Note that implementing a quantizer $Q : (\mathcal{X}, \mathcal{T}) \to \mathcal{Z}$ can be simplified to computing a set of quantizers corresponding to each $\theta \in \mathcal{T}$ as in Fig. 2 without loss of generality. If the quantizer does not include a region m for some realization of θ, the encoder never sends the message m i.e., the encoder chooses a lower rate and is less revealing for that value of θ. In Fig. 2, we see that the quantizer q_{θ_1} only includes $m = 1, 2, 4$ regions, while the quantizer q_{θ_2} contains all five regions.

III. Main Results

In this section, we first present the gradient-descent based optimization algorithm. We also present a straightforward method of quantization along with linear estimation that results in an upper bound on the encoder distortion.

A. Proposed algorithm

In [11], we proposed a gradient-descent based algorithm to solve the problem of quantization of a scalar source with misaligned encoder and decoder objectives communicating over a fixed rate noiseless channel. We extend this algorithm to a 2-dimensional source (X, θ) by a simple method of computing quantizers for each value of θ as $Q = \{q_\theta, \theta \in \mathcal{T}\}$. The gradient descent optimization is performed with the objective as the encoder distortion optimized over the encoder's choice of quantizer decision levels $Q = \{q_\theta, \theta \in \mathcal{T}\}$. Although the encoder distortion depends on decoder reconstruction levels $\mathbf{y}$, since $\mathbf{y}$ is a function of Q, the optimization can be implemented as a function of solely Q.

Remark 2. *The proposed method inherits the convergence guarantees of gradient-descent based algorithms. Hence local optimality is guaranteed, but the resulting quantizer may not necessarily be globally optimal.*

Algorithm 1 Proposed strategic quantizer design

Parameters: ϵ, λ
Input: $f(\cdot,\cdot), \mathcal{X}, \mathcal{T}, M, \eta_E, \eta_D$
Output: $\{q_\theta^*\}$, $\{y_m^*\}$, D_E, D_D
Initialization: assign a set of monotone $\{q_{\theta,0}\}$ randomly, compute associated encoder distortion $D_E(0)$, set iteration index $i = 1$;
while $\Delta D > \epsilon$ *or until a set amount of iterations* **do**
  compute the gradients $\{\partial D_E / \partial x_{\theta,:}\}_i$,
  compute the updated quantizer $q_{\theta,i+1} \triangleq q_{\theta,i} - \lambda\{\partial D_E / \partial x_{\theta,:}\}_i$ for $\theta \in \mathcal{T}$,
  compute actions $\mathbf{y}(\{q_{\theta,i+1}\})$ via (2),
  compute encoder distortion $D_E(i+1)$ associated with quantizer values $q_{\theta,i+1}$ and actions $\mathbf{y}(\{q_{\theta,i+1}\})$ via (1),
  compute $\Delta D = D_E(i) - D_E(i+1)$.
return quantizer $\{q_\theta^*\} = \{q_{\theta,i+1}\}$, actions $\{y_m^*\} = \mathbf{y}(\{q_\theta^*\})$, encoder and decoder distortions D_E and D_D computed for optimal quantizer and decoder actions $\{q_\theta^*\}, \mathbf{y}(\{q_\theta^*\})$ via (1).

Like any gradient-descent-based algorithm, the proposed method may get stuck at a poor local optimum, which can be resolved with different techniques [19]–[21]. As a simple remedy, we perform gradient descent with multiple initializations and choose the best local optimum among them. A sketch of the proposed method is summarized in Algorithm 1. The MATLAB codes are provided at `https://tinyurl.com/allerton24` for research purposes.

Note that a strategic variation of Lloyd-Max optimization may not result in a local optima, as shown in detail in [11].

B. Upper bound

Conventional compression problems with quadratic (MSE) distortion and an additive noisy source admit a decomposition where the overall distortion can be expressed as a sum of estimation distortion and compression distortion, see e.g. [22], [23]. The fact that our problem $\eta_E(X,\theta,\hat{X}) = (X+\theta-\hat{X})^2$ resembles an additive noisy source inspired us to investigate whether similar techniques can be used to obtain closed-form solutions for special cases (such as jointly Gaussian sources).

Moreover, the information-theoretic solution to the strategic compression problem with quadratic measures (as analyzed in [4]) admits the following solution for jointly Gaussian sources $(X,\theta) \sim \mathcal{N}\left(\begin{bmatrix}0\\0\end{bmatrix}, \sigma_X^2\begin{bmatrix}1 & \rho\\ \rho & r\end{bmatrix}\right), r > \rho^2$: the encoder compresses, in rate-distortion sense, the linear combination $T = X + \alpha\theta$, where $\alpha = (A-1)/(2(r+\rho))$, $A = \sqrt{1+4(r+\rho)}$, and the decoder reconstruction is

$$\hat{X} = \mathbb{E}\{X|T\} = \kappa T, \quad T = X + \alpha\theta,$$

where $\kappa = (1+\alpha\rho)/(1+\alpha^2 r + 2\alpha\rho)$.

Inspired by these observations, we now propose a scheme similar to the information-theoretic solution described above for jointly Gaussian sources. The method proposed below computes an upper bound for the encoder distortion for (X,θ) following a general distribution (not necessarily jointly Gaussian), and is based on linear estimation strategies.

We consider the linear minimum mean squared error (LMMSE) estimate of X given observation T,

$$\hat{X} = LMMSE(X|T) = h(T) = \kappa T, \tag{3}$$

$$T = g(X,\theta) = (X+\alpha\theta). \tag{4}$$

where the parameters α and κ are

$$\alpha = \frac{A-1}{2(r+\rho)}, \quad \kappa = \frac{1+\alpha\rho}{1+\alpha^2 r + 2\alpha\rho}, \tag{5}$$

and A, r, and ρ are given by

$$A = \sqrt{1+4(r+\rho)}, \quad r = \frac{\sigma_\theta^2}{\sigma_X^2}, \quad \rho = \frac{\mathbb{E}\{X\theta\}}{\sigma_X^2}. \tag{6}$$

The encoder distortion D_E can be written as,

$$\begin{aligned} D_E &= \mathbb{E}\{(X+\theta-Q(\hat{X}(T)))^2\} \\ &= \mathbb{E}\{(X+\theta-\hat{X}(T)+\hat{X}(T)-Q(\hat{X}(T)))^2\} \\ &\overset{a}{=} \mathbb{E}\{(X+\theta-\hat{X}(T))^2\} + \mathbb{E}\{(\hat{X}(T)-Q(\hat{X}(T)))^2\} \\ &\quad + 2\mathbb{E}\{\theta(\hat{X}(T)-Q(\hat{X}(T)))\}. \end{aligned} \tag{7}$$

Equality a in the above equation is due to the orthogonality of the estimation error $(X-\hat{X}(T))$ to any function of the observation T,

$$\mathbb{E}\{(X-\hat{X}(T))(\hat{X}(T)-Q(\hat{X}(T)))\} = 0.$$

Remark 3. *Similar decompositions were also used in [22], [23], where they exploit the orthogonality of the estimation error.*

Let us define the optimal quantizer that minimizes (1) as Q^*,

$$\begin{aligned} Q^* &= \arg\min_Q D_E \\ &= \arg\min_Q \Big\{\mathbb{E}\{(X+\theta-\hat{X}(T))^2\} + 2\mathbb{E}\{\theta\hat{X}(T)\} \\ &\quad + \mathbb{E}\{(\hat{X}(T)-Q(\hat{X}(T)))^2\} - 2\mathbb{E}\{\theta Q(\hat{X}(T))\}\Big\} \\ &\overset{b}{=} \arg\min_Q \Big\{\mathbb{E}\{(\hat{X}-Q(\hat{X}))^2\} - 2\mathbb{E}\{\theta Q(\hat{X})\}\Big\} \end{aligned} \tag{8}$$

where equality b is due to the fact that the first two terms $\mathbb{E}\{(X+\theta-\hat{X}(T))^2\}$ and $2\mathbb{E}\{\theta\hat{X}(T)\}$ do not include Q.

In general, it is hard to compute Q^*. Instead, we consider Q^{**} which we define as

$$Q^{**} = \arg\min_Q \mathbb{E}\{(\hat{X}-Q(\hat{X}))^2\}. \tag{9}$$

In other words, Q^{**} is the mean squared error (MSE) optimal non-strategic quantizer for $\hat{X} = LMMSE(X|T) = \kappa T$, $T = (X+\alpha\theta)$.

We note that Q^{**} is relatively straightforward to compute, e.g., if (X,θ) is jointly Gaussian, $\hat{X}$ is also Gaussian for which the optimal quantizer is well-known, e.g., [24].

Since $Q^{**} \neq Q^*$ in general, the resulting distortion of Q^{**}, denoted by D_E^U is an upper bound, i.e., $D_E^U \geq D_E$.

We formalize the preceding discussion in the following theorem:

Theorem 2. *$D_E^U \geq D_E$, where*

$$\begin{aligned} D_E^U &= \mathbb{E}\{(X+\theta-\hat{X}(T))^2\} + \mathbb{E}\{(\hat{X}-Q^{**}(\hat{X}))^2\} \\ &\quad + 2\mathbb{E}\{\theta(\hat{X}-Q^{**}(\hat{X}))\}, \end{aligned}$$

and $T = X+\alpha\theta$, $\hat{X} = LMMSE(X|T) = \kappa T$,

$$\alpha = \frac{A-1}{2(r+\rho)}, \quad \kappa = \frac{1+\alpha\rho}{1+\alpha^2 r+2\alpha\rho}$$

$$r = \frac{\sigma_\theta^2}{\sigma_X^2}, \quad \rho = \frac{\mathbb{E}\{X\theta\}}{\sigma_X^2}, \quad A = \sqrt{1+4(r+\rho)},$$

and Q^{**} is given in (9).

Proof: Since the set of quantizers over which the encoder optimizes its distortion also includes this specific scheme involving linear strategies which may not be the optimal quantizer, D_E^U is an upper bound to D_E.

In summary, for the computation of an upper bound, we consider a system where the encoder computes $\hat{X} = h(g(X,\theta)) = \kappa(X+\alpha\theta)$, quantizes $\hat{X}$ as $Z = Q^{**}(\hat{X})$ and

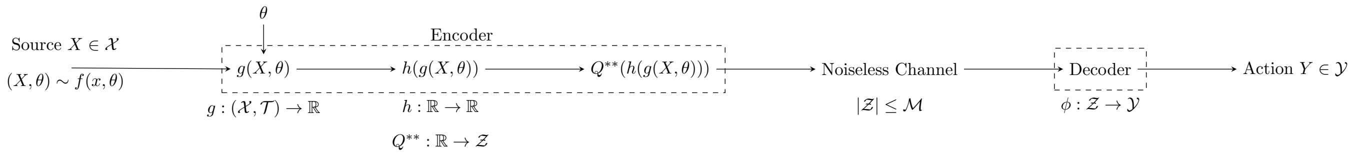

Fig. 3. Computation of an upper bound for the encoder distortion.

sends the message Z to the decoder, as depicted in Fig. 3. The upper bound is computed as the sum of the estimation error, quantization error, and the term $2\mathbb{E}\{\theta(\hat{X}-Q^{**}(\hat{X}))\}$, as in (7). A sketch of the computation of this upper bound is in Algorithm 2 below.

Algorithm 2 Computation of an upper bound of encoder distortion

Input: $f(\cdot,\cdot), \mathcal{X}, \mathcal{T}, M, \eta_E, \eta_D$
Output: Q^{**}, D_E^U
Compute r, ρ, A from (6).
$\alpha \leftarrow (A-1)/(2(r+\rho))$
$\kappa \leftarrow (1+\alpha\rho)/(1+\alpha^2 r+2\alpha\rho)$
Compute probability distribution function of $\hat{X}=\kappa(X+\alpha\theta)$, $f_{\hat{X}}$.
Compute non-strategic quantizer Q^{**} with MSE encoder and decoder objectives $\mathbb{E}\{(\hat{X}-Q^{**}(\hat{X}))^2\}$, and $\hat{X}\sim f_{\hat{X}}$.
Compute the upper bound as D_E^U in Theorem 2.
return quantizer Q^{**}, upper bound D_E^U

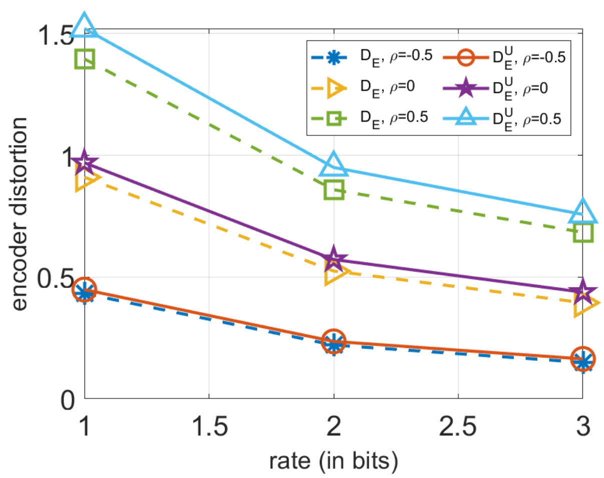

Fig. 4. Encoder distortion and the associated upper bound for a jointly Gaussian source $(X,\theta)\sim\mathcal{N}\left(\begin{bmatrix}0\\0\end{bmatrix},\begin{bmatrix}1 & \rho\\ \rho & 1\end{bmatrix}\right)$ with $\eta_E(x,\theta,y)=(x+\theta-y)^2, \eta_D(x,\theta,y)=(x-y)^2$.

IV. Numerical Results

We present results for two settings with encoder and decoder distortions $\eta_E(x,\theta,y)=(x+\theta-y)^2$, $\eta_D(x,y)=(x-y)^2$:

- in Fig. 4, we present results for a jointly Gaussian source $(X,\theta)\sim\mathcal{N}\left(\begin{bmatrix}0\\0\end{bmatrix},\begin{bmatrix}1 & \rho\\ \rho & 1\end{bmatrix}\right)$ for $\rho\in\{-0.5,0,0.5\}$,
- in Fig. 5, we present results for independent (X,θ), Uniform $X\sim U[0,1]$, Bernoulli $\theta\in\{-1,1\}$ with probability $[0.5,0.5]$.

The support of θ is discretized by sampling for computational feasibility for the jointly Gaussian source.

The encoder and decoder distortions for the two settings are shown in Figures 4 and 5 respectively. As expected from Theorem 2, we observe that $D_E^U\geq D_E$.

The numerical results suggest that the upper bound is tighter for jointly Gaussian settings. This is also expected since LMMSE used in computing the upper bound coincides with MMSE estimator for a jointly Gaussian distribution. Moreover, the upper bound seems to get tighter as correlation decreases. We leave the theoretical analysis of such observations as future work.

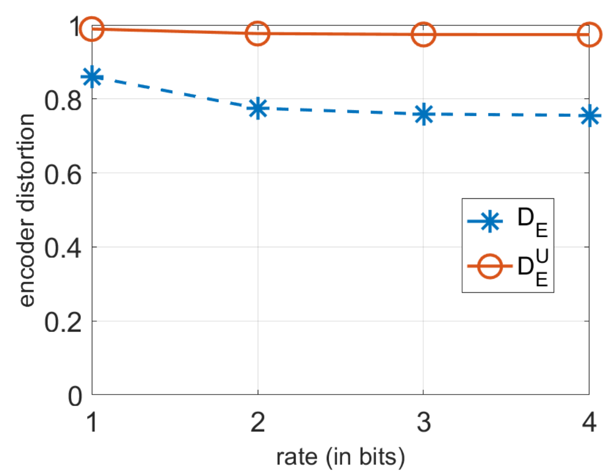

Fig. 5. Encoder distortion and the associated upper bound for $X\sim U[0,1]$, and a Bernoulli source $\theta\in[-1,1]$ with probability $p_\theta=[0.5,0.5]$, X,θ independent $\eta_E(x,\theta,y)=(x+\theta-y)^2, \eta_D(x,\theta,y)=(x-y)^2$.

V. Conclusions

In this paper, we analyzed the problem of strategic quantization of a 2-dimensional source (X,θ) with the encoder and the decoder objectives $D_E=\mathbb{E}\{(X+\theta-Y)^2\}$ and $D_D=\mathbb{E}\{(X-Y)^2\}$, respectively. We extended our prior work on design, i.e, a gradient-descent based algorithm for scalar sources, to the setting considered in this paper. We finally presented a method to compute an upper bound for the encoder distortion based on linear communication strategies. The numerical results obtained via the proposed algorithm and upper bound suggest several intriguing research problems which we leave as a part of our future work.

REFERENCES

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

[2] S. Dughmi and H. Xu, "Algorithmic Bayesian Persuasion," *SIAM Journal on Computing*, vol. 50, no. 3, pp. STOC16–68–STOC16–97, 2021. [Online]. Available: https://doi.org/10.1137/16M1098334

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

[4] E. Akyol, C. Langbort, and T. Başar, "Information-Theoretic Approach to Strategic Communication as a Hierarchical Game," *Proceedings of the IEEE*, vol. 105, no. 2, pp. 205–218, 2016.

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

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

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

[8] M. Le Treust and T. Tomala, "Persuasion with Limited Communication Capacity," *Journal of Economic Theory*, vol. 184, no. C, 2019. [Online]. Available: https://ideas.repec.org/a/eee/jetheo/v184y2019ics0022053118305064.html

[9] R. Deori and A. A. Kulkarni, "Information Revelation Through Signalling," 2022. [Online]. Available: https://arxiv.org/abs/2202.10145

[10] A. Anand and E. Akyol, "Optimal Strategic Quantizer Design via Dynamic Programming," in *Proceedings of the IEEE Data Compression Conference*. IEEE, 2022, pp. 173–181.

[11] E. Akyol and A. Anand, "Strategic Quantization," in *2023 IEEE International Symposium on Information Theory (ISIT)*, 2023, pp. 543–548.

[12] A. Anand and E. Akyol, "Channel-Optimized Strategic Quantizer Design via Dynamic Programming," in *2023 IEEE Statistical Signal Processing Workshop (SSP)*, 2023, pp. 621–625.

[13] ——, "Strategic Quantization over a Noisy Channel," in *2023 57th Asilomar Conference on Signals, Systems, and Computers*, 2023. Available online, at arXiv.

[14] ——, "Strategic Quantization of a Noisy Source," in *2023 59th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, 2023, pp. 1–7.

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

[16] P. E. Fischer and R. E. Verrecchia, "Reporting Bias," *The Accounting Review*, vol. 75, no. 2, pp. 229–245, 2000.

[17] R. Bénabou and J. Tirole, "Incentives and Prosocial Behavior," *American Economic Journal: Microeconomics*, vol. 95, no. 5, pp. 1652–1678, 2006.

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

[19] K. Rose, "Deterministic Annealing for Clustering, Compression, Classification, Regression, and Related Optimization Problems," *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2210–2239, 1998.

[20] D. Bertsimas and J. Tsitsiklis, "Simulated Annealing," *Statistical Science*, vol. 8, no. 1, pp. 10 – 15, 1993. [Online]. Available: https://doi.org/10.1214/ss/1177011077

[21] S. Gadkari and K. Rose, "Robust Vector Quantizer Design by Noisy Channel Relaxation," *IEEE Transactions on Communications*, vol. 47, no. 8, pp. 1113–1116, 1999.

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

[23] E. Ayanoglu, "On Optimal Quantization of Noisy Sources," *IEEE Transactions on Information Theory*, vol. 36, no. 6, pp. 1450–1452, 1990.

[24] J. Max, "Quantizing for Minimum Distortion," *IRE Transactions on Information Theory*, vol. 6, no. 1, pp. 7–12, 1960.