On the Impact of Bounded Rationality in Strategic Data Gathering ⋆

Anju Anand * **Emrah Akyol** **

* *Binghamton University–SUNY, Binghamton, NY, 13902 USA (e-mail: aanand6@binghamton.edu)*
** *Binghamton University–SUNY, Binghamton, NY, 13902 USA (e-mail: eakyol@binghamton.edu)*

Abstract: We consider the problem of estimation from survey data gathered from strategic and boundedly-rational agents with heterogeneous objectives and available information. Particularly, we consider a setting where there are three different types of survey responders with varying levels of available information, strategicness, and cognitive hierarchy: i) a non-strategic agent with an honest response, ii) a strategic agent that believes everyone else is a non-strategic agent and that the decoder also believes the same, hence assumes a naive estimator, i.e., level-1 in cognitive hierarchy iii) and strategic agent that believes the population is Poisson distributed over the previous types, and that the decoder believes the same. We model each of these scenarios as a strategic classification of a 2-dimensional source (possibly correlated source and bias components) with quadratic distortion measures and provide a design algorithm. Finally, we provide our numerical results and the code to obtain them for research purposes at `https://tinyurl.com/CPHS2024-bounded-rationality`.

Keywords: Behavioral sciences, Game theory, Bayesian estimation, Mathematical models, Human behavior modeling

1 Introduction

Consider the following scenario involving a survey designed to gauge public reception of a new plastic product, with responses influenced by respondents' attitudes toward climate change. Respondents' scores range from 1 ('will definitely not use') to 4 ('will definitely use'), and the survey needs to account for potential biases as well as varying levels of rationality among respondents. We model this problem using the hierarchical cognitive type model as studied by Camerer et al. (2004), considering three types of respondents:

- Type 0 (Honest-Nonstrategic Respondents): These respondents provide truthful information based on their actual opinions about the product, unaffected by their considerations of climate change or any desire to bias the survey.
- Type 1: These respondents wish to influence the survey outcome correlated with their attitudes. They best respond to Type 0, assuming that
  (1) All other respondents are of Type 0.
  (2) The estimator (designer) is only aware of Type 0 respondents.
- Type 2: These respondents have a higher level of strategic thinking and behave as the best response to a mix of Types 0 and 1, assuming that the designer (estimator) perceives the responses as coming from a distribution of these lower types.

The designer of the survey is aware of the existence of these types of respondents as well as their true statistics. The question explored in this paper is: What is the designer's optimal "de-biasing" procedure, i.e, optimally (in Bayesian sense) estimating the unbiased scores that reflect the true public reception of the plastic product?

We approach this problem via the recently introduced strategic quantization framework, see Akyol and Anand (2023), which is a special case of the information design problem in Economics.[1] This class of problems, notable studied by Kamenica and Gentzkow (2011); Rayo and Segal (2010) explore the use of information by an agent (sender) to influence the action taken by another agent (receiver), where the aforementioned action determines the payoffs for both agents. Our prior work explored strategic quantization problem settings where the sender and the receiver were assumed to be fully rational agents. In this paper, we extend our strategic quantization work to settings with boundedly-rational sender (quantizer), via employing the cognitive hierarchy model of Camerer et al. (2004).

Throughout this paper, we focus on the quadratic distortion measures. Particularly, the senders observe a two-dimensional source $(X, S) \sim f_{X,S}(\cdot, \cdot)$ with a known joint density function over X and S, where X and S can be interpreted as the state and bias variables. There are two types of strategic senders, both trying to minimize $\{(X + S - \hat{X})^2\}$, with different assumptions on the estimator

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

[1] Throughout the paper, we use the terms quantizer and classifier interchangeably.

(receiver). Type 1 strategic users assume the estimator is simply nonstrategic (which is the best response to Type 0). Type 2, in turn, assumes that the decoder is aware of a mix of Type 0 and Type 1 senders.

The receiver's objective is to estimate the true state in the minimum mean squared error (MME) sense, i.e., the receiver minimizes $\eta(x,y) = (x-y)^2$ by choosing an action $\hat{X}$ which is the optimal MMSE estimate of x given the quantization index from the sender $y = Q(x,s)$, hence $\hat{X} = \mathbb{E}\{X|Y=y\}$. In sharp contrast with the conventional quantization problem where the sender chooses Q that minimizes $\mathbb{E}\{(X-\hat{X})^2\}$, in this setting the sender's choice of quantization mapping Q minimizes a biased estimate, i.e., $\mathbb{E}\{(X+S-\hat{X})^2\}$. The objectives and the source distribution are common knowledge, available for all agents. We note that similar signaling problems with quadratic measures have been analyzed in the Economics literature, see e.g., Bénabou and Tirole (2006); Crawford and Sobel (1982); Fischer and Verrecchia (2000).

This paper is organized as follows: In Section 2 we present the problem formulation. In Section 3, we present a gradient-descent based algorithm to compute the classifier implemented by the boundedly rational agent. We provide numerical results in Section 4, and conclude in Section 5.

2 Preliminaries

2.1 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 ($\mathcal{X}$). Vectors are denoted in bold font. 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 2-dimensional jointly Gaussian probability density function with mean $[t_1 \;\; t_2]'$ and respective variances σ_1^2, σ_2^2 with a correlation ρ is denoted by $\mathcal{N}\left(\begin{bmatrix} t_1 \\ t_2 \end{bmatrix}, \begin{bmatrix} \sigma_1^2 & \sigma_1\sigma_2\rho \\ \sigma_1\sigma_2\rho & \sigma_2^2 \end{bmatrix}\right), 0 \le \rho < 1,\ t_1, t_2 \in \mathbb{R}$. 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.

2.2 Problem Formulation

Consider the following classification problem: Three classifiers (senders), $E_k, k \in [0:2]$ each with a probability p_k of being chosen to send the message observe realizations of the two sources $X \in \mathcal{X} \subseteq [a_X, b_X], S \in \mathcal{S} \subseteq [a_S, b_S]$, $a_X, b_X, a_S, b_S \in \mathbb{R}$ with joint probability density $(X,S) \sim f_{X,S}(\cdot,\cdot)$. The chosen classifier E_k maps (X,S) to a message $Z \in \mathcal{Z}$, where $\mathcal{Z}$ is a set of discrete messages with a cardinality constraint $|\mathcal{Z}| \le M$ using a non-injective mapping $Q^k : (\mathcal{X}\times\mathcal{S}) \to \mathcal{Z}$. After receiving the message Z, the receiver applies a mapping $\phi : \mathcal{Z} \to \mathcal{Y}$ on the message Z and takes an action $Y = \phi(Z)$.

The set $\mathcal{X}$ is divided into mutually exclusive and exhaustive sets by each classifier E_k as $\mathcal{V}_1^k, \mathcal{V}_2^k, \dots, \mathcal{V}_M^k$. Let the marginal probability density function of X be $f_X(x)$.

The probability p_k of sender E_k being chosen follows a normalized Poisson distribution

$$p_k = \frac{e^{\lambda}\frac{\lambda^k}{k!}}{\sum_{i=0}^{2} e^{\lambda}\frac{\lambda^i}{i!}}. \quad (1)$$

The distortion of the senders E_1, E_2 are $\eta_E^k(x,s,y) = (x+s-y)^2, k \in [1:2]$, that of sender E_0 is $\eta_E^0(x,s,y) = (x-y)^2$, and that of the receiver is $\eta_D(x,y) = (x-y)^2$. Let $y_m^{(k)}$ be the estimates that E_k assumes are optimized by the receiver with respect to the respective perceived receiver objectives D_D^k, obtained by enforcing KKT conditions of optimality, $\partial D_D^k/\partial y_m^{(k)} = 0$. We consider three senders with hierarchical cognitive types and define the senders' and their respective perceived receiver objectives as $D_E^k, D_D^k, k \in [0:2]$ below:

(1) Non-strategic sender E_0: similar to level L_0 cognitive type, the sender assumes all senders are of type E_0, and that the decoder assumes all senders are of type E_0. Sender E_0 considers the receiver's objective as the same as the sender's, $D_E^0 = D_D^0$ (provides the information required by the receiver honestly)
$$D_D^0 = \sum_{m=1}^{M} \mathbb{E}_S\{\eta_D(X,S,y_m^{(0)})|X \in \mathcal{V}_m^0\}.$$
(2) Level-1 strategic sender E_1: similar to level L_1 cognitive type, the sender assumes all other senders are of type E_0 and that it is uniquely of type E_1. The sender assumes the receiver thinks that all sender types are E_0, i.e., $D_D^1 = D_D^0$, which results in the estimates perceived by E_1, $\mathbf{y}^{(1)} = \mathbf{y}^{(0)}$.
(3) Level-2 strategic sender E_2: similar to level L_2 cognitive type, the sender assumes the other senders are of lower cognitive levels and are Poisson distributed with a probability mass function $\mathbf{p}' = [p_0', p_1']$,
$$p_k' = \frac{e^{\lambda}\frac{\lambda^k}{k!}}{\sum_{i=0}^{1} e^{\lambda}\frac{\lambda^i}{i!}},$$
for $E_k, k \in [0:1]$, respectively. Note that this perceived probability mass function $\mathbf{p}'$ is not the actual statistics of the population, $\mathbf{p}' \neq \mathbf{p}$. The sender assumes the receiver is aware only of the types E_0 and E_1 and its perceived probability mass function p_0' and p_1'. Sender E_2's perceived receiver objective,
$$D_D^2 = \sum_{i=0}^{1} p_i' \sum_{m=1}^{M} \mathbb{E}_S\{\eta_D(X,S,y_m^{(2)})|X \in \mathcal{V}_m^i\}.$$

The encoder distortions for each type $k \in [0:2]$,

$$D_E^k = \sum_{m=1}^{M} \mathbb{E}_S\{\eta_E^k(X,S,y_m^{(k)})|X \in \mathcal{V}_m^k\}.$$

The receiver objective is given by

$$D_D^* = \min_{\mathbf{y}^*} \sum_{i=0}^{2} \sum_{m=1}^{M} p_i \mathbb{E}_S\{\eta_t^i(X,S,y_m^*|X \in \mathcal{V}_m^i\}.$$

Each sender type $E_k, k \in [1:2]$ optimizes their classifiers Q^k with respect to their own objective D_E^k, assuming the receiver is aware of only $E_i, i < k$ sender types. Sender $E_k, k \in [0:2]$ designs Q^k *ex-ante*, i.e., without

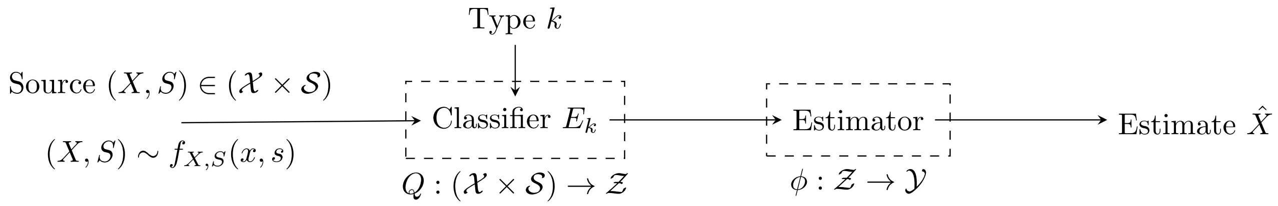


Fig. 1. Communication diagram: with probability p_k, E_k sends a message Z which is a function of the source (X,S) over a noiseless channel

the knowledge of the realization of (X,S), using only the objectives D_E^k and D_D^k, and the statistics of the source $f_{X,S}(\cdot,\cdot)$.

The receiver is fully rational and has full information about the classification setup. The shared prior ($f_{X,S}$), the probability mass function over the sender types ($\mathbf{p} = [p_0, p_1, p_2]$) and the mappings ($\mathbf{Q} = \{Q^k, k \in [0:2]\}$) are known to the receiver. The problem is to design the classifiers $\mathbf{Q}$ for the equilibrium, i.e., each sender type E_k minimizes its own objective, assuming that the receiver minimizes its corresponding perceived objective D_D^k. This classification problem is given in Fig. 1. Since the senders choose the classifiers $\mathbf{Q}$ first, followed by the receiver choosing the perceived estimates ($\mathbf{y}^{(k)}, k \in [0:2]$), we look for a Stackelberg equilibrium.

The classifier design involves computing classifiers for each realization of S by classifier i as $\mathcal{U}_{s,m}^i, s \in \mathcal{S}$, where $\bigcup_{s \in \mathcal{S}} \mathcal{U}_{s,m}^i = \mathcal{V}_m^i$. Throughout this paper, we make the following "monotonicity" assumption on the sets $\{U_{s,m}^i\}$.

Assumption 1. *$U_{s,m}^i$ is convex for all $m \in [1:M], s \in \mathcal{S}, i \in [0:2]$.*

That is, $\mathcal{U}_{s,m}^i = [q_{s,m}^i, q_{s,m+1}^i], q_{s,m}^i < q_{s,m+1}^i$. Then, $Q^i = \{\mathbf{q}_s^i, s \in \mathcal{S}\}$, where $\mathbf{q}_s^i = [q_{s,0}^i, \ldots, q_{s,M+1}^i]$.

Since E_0's distortion function $\eta_E^0(x,s,y) = (x-y)^2$ is not a function of s, $Q^0 : (\mathcal{X} \times \mathcal{S}) \to \mathcal{Z}$ simplifies to $Q^0 : \mathcal{X} \to \mathcal{Z}$. Let $\mathbf{q}_s^0 = \mathbf{n} = [n_0, \ldots, n_M]$ for all $s \in \mathcal{S}$. E_0 responds honestly, $D_D^0 = D_E^0$ (equivalent to the non-strategic classification setting), hence its classifier $\mathbf{n}$ and perceived estimates $\mathbf{y}^{(0)}$ are,

$$D_E^0 = \sum_{m=1}^{M} \int_{n_{m-1}}^{n_m} (x - y_m^{(0)})^2 f_X(x) \mathrm{d}x, \quad y_m^{(0)} = \mathbb{E}\{X | x \in \mathcal{V}_m^0\}.$$

E_1 assumes that all other senders are of type E_0 and that the receiver views all senders as type E_0, i.e., E_1's perceived estimates $y_m^{(1)} = y_m^{(0)}$, since $D_D^1 = D_D^0$.

E_2 assumes the other agents are of types E_0 and E_1 with probability mass function $\mathbf{p}'$, and that the receiver perceives the agents as the same, of types E_0, E_1 with a probability mass function $\mathbf{p}'$,

$$D_D^2 = \sum_{i=0}^{1} p_i' \sum_{m=1}^{M} \int_{a_S}^{b_S} \int_{q_{s,m-1}^i}^{q_{s,m}^i} (x - y_m^{(2)})^2 f_{X,S}(x,s) \mathrm{d}x \mathrm{d}s,$$

resulting in E_2's perceived estimates $y_m^{(2)}$,

$$y_m^{(2)} = \frac{\sum_{i=0}^{1} p_i' \int_{a_S}^{b_S} \int_{q_{s,m-1}^i}^{q_{s,m}^i} x f_{X,S}(x,s) \mathrm{d}x \mathrm{d}s}{\sum_{i=0}^{1} p_i' \int_{a_S}^{b_S} \int_{q_{s,m-1}^i}^{q_{s,m}^i} f_{X,S}(x,s) \mathrm{d}x \mathrm{d}s}. \tag{2}$$

The classifier objectives for $E_k, k \in [1:2]$ are given by

$$D_E^k = \sum_{m=1}^{M} \int_{a_S}^{b_S} \int_{q_{s,m-1}^k}^{q_{s,m}^k} (x + s - y_m^{(k)})^2 f_{X,S}(x,s) \mathrm{d}x \mathrm{d}s. \tag{3}$$

The receiver's objective and estimates are

$$D_D^* = \sum_{i=0}^{2} p_i \sum_{m=1}^{M} \int_{a_S}^{b_S} \int_{q_{s,m-1}^i}^{q_{s,m}^i} (x - y_m^*)^2 f_{X,S}(x,s) \mathrm{d}x \mathrm{d}s, \tag{4}$$

$$y_m^* = \frac{\sum_{i=0}^{2} p_i \int_{a_S}^{b_S} \int_{q_{s,m-1}^i}^{q_{s,m}^i} x f_{X,S}(x,s) \mathrm{d}x \mathrm{d}s}{\sum_{i=0}^{2} p_i \int_{a_S}^{b_S} \int_{q_{s,m-1}^i}^{q_{s,m}^i} f_{X,S}(x,s) \mathrm{d}x \mathrm{d}s}. \tag{5}$$

The classifers implemented by E_0, E_1, E_2 are as follows:

(1) E_0 implements a non-strategic (classical) classifier $\mathbf{n}$ for the given density of X, $f_X(\cdot)$.
(2) E_1 implements a non-strategic classifier $\mathbf{q}_s^1$ for $X + s$ with the probability density function $f_X(x-s)$ for $s \in \mathcal{S}$. Since E_1 assumes that the receiver views all senders as type E_0, the estimates perceived by E_1 is the non-strategic estimates $y_m^{(1)} = y_m^{(0)}$. Minimizing D_E^1 with $y_m^{(0)}$ estimates results in a nearest neighbor classifier, which is the non-strategic classifier shifted by s for each realization $s \in \mathcal{S}$, $\mathbf{q}_s^1 = \mathbf{n} + s$.
(3) E_2 implements the classifier minimizing (3) with the given source probability density function $f_{X,S}(x,s)$ assuming receiver estimates are given by (2).

3 Design

In this section, we present our gradient-descent based algorithm for the optimization of Q^2.

In Anand and Akyol (2024), we proposed a gradient-descent based algorithm to solve the problem of quantization of a 2-dimensional source (X,S) by extending our algorithm in Akyol and Anand (2023) for a scalar

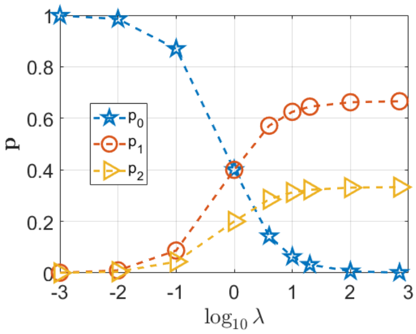


Fig. 2. Probability mass function $\mathbf{p}$ with respect to λ.

source to the 2-dimensional setting by a simple method of computing quantizers for each value of $s \in \mathcal{S}$.

Here, we use similar methods as in Anand and Akyol (2024), also using the known classifiers for E_0, E_1, we perform gradient descent optimization with the objective as D_E^2 optimized over Q^2, assuming the estimates $\mathbf{y}^{(2)}$ are optimized for D_D^2. Although the sender's objective D_E^2 depends on receiver estimates $\mathbf{y}^{(2)}$, since $\mathbf{y}^{(2)}$ is a function of Q^2, the optimization can be implemented as a function of solely Q^2.

Algorithm 1 Proposed strategic quantizer design

Parameters: ϵ, λ
Input: $f_{X,S}(\cdot,\cdot), \mathcal{X}, \mathcal{S}, M, \eta_E^2, \eta_D^2, Q^0, Q^1, \mathbf{p}, \mathbf{p}'$
Output: Q^2, $\mathbf{y}^{(2)}$, $\mathbf{y}^*$, D_E^2, D_D^*
Initialization: assign a set of monotone $\{\mathbf{q}_s^2\}_0$ randomly, compute associated sender distortion $D_E^2(0)$, set iteration index $j = 1$;
while $\Delta D > \epsilon$ *or until a set amount of iterations* **do**
  compute the gradients $\{\partial D_E^2/\partial \mathbf{q}_s^2\}_j$,
  compute the updated classifier $\{\mathbf{q}_s^2\}_{j+1} \triangleq \{\mathbf{q}_s^2\}_j - \lambda\{\partial D_E^2/\partial \mathbf{q}_s^2\}_j$ for $s \in \mathcal{S}$,
  compute estimates $\mathbf{y}(\{\mathbf{q}_s^2\}_{j+1})$ via (2),
  compute sender objective $D_E^2(j+1)$ associated with classifier $\{\mathbf{q}_s^2\}_{j+1}$ and estimates $\mathbf{y}(\{\mathbf{q}_s^2\}_{j+1})$ via (3),
  compute $\Delta D = D_E^2(j) - D_E^2(j+1)$.
return classifier $Q^2 = \{\mathbf{q}_s^2\}_{j+1}$, E_2's perceived receiver estimates $\mathbf{y}^{(2)} = \mathbf{y}(Q^2)$, actual receiver estimates $\mathbf{y}^*$, sender objective D_E^2 computed for the bounded rational optimal classifiers $\mathbf{Q} = \{Q^k, \mathbf{y}^{(k)}, k \in [0,2]\}$ with the perceived receiver estimates $\mathbf{y}^{(2)}$, and receiver objective D_D^* computed for the bounded rational optimal classifiers $\mathbf{Q}$ with the actual receiver estimate $\mathbf{y}^*$ via (5), (3), (4).

Like any gradient-descent-based algorithm, the proposed method may get stuck at a poor local optimum, which we resolve with a simple remedy by performing gradient descent with multiple initializations and choosing 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/CPHS2024-bounded-rationality` for research purposes.

4 Numerical Results

We consider a jointly Gaussian 2-dimensional source

$$(X,S) \sim \mathcal{N}\left(\begin{bmatrix}0\\0\end{bmatrix}, \begin{bmatrix}1 & \sigma_S\rho\\ \sigma_S\rho & \sigma_S^2\end{bmatrix}\right)$$

and present results for different settings with parameters bias variance $\sigma_S^2 \in \{0.1, 1, 1.5\}$, correlation $\rho \in \{0.1, 0.5, 0.7\}$, and $\mathbf{p}$ following (1) with $\lambda \in [0.001, 700]$ for an $M = 4$ classifier. The probability mass function over the sender types $\mathbf{p}$ for different values of λ is plotted in Fig. 2. We consider only positive correlation since people's preferences are positively correlated with their opinions; for instance, both climate activists and climate change deniers try to bias their classification towards the extremes on their side. We plot the receiver's objective D_D^* in Fig. 3 for given values of σ_S^2 and ρ, respectively.

We now interpret our results in terms of the impact of different parameters on the receiver's estimation.

We observe from Fig. 3 that the correlation between the state and bias variables ρ does not change the receiver objective trends.

4.1 Impact of cognitive parameter λ

From Fig. 2 we note that as $\lambda \to 0$, the population mostly consists of level-0 cognitive level. As λ increases, the population shifts towards higher cognitive types, and we expect the receiver objective to increase with λ, as we observe in Fig. 3. For $\lambda \to 0$, the receiver objective does not change significantly with varying bias variance σ_S^2 since the population is mostly of level-0 type, and they respond honestly. For $\lambda > 100$, the statistics of the population remain fairly constant, and hence the receiver objective varies negligibly.

4.2 Impact of varying σ_S^2

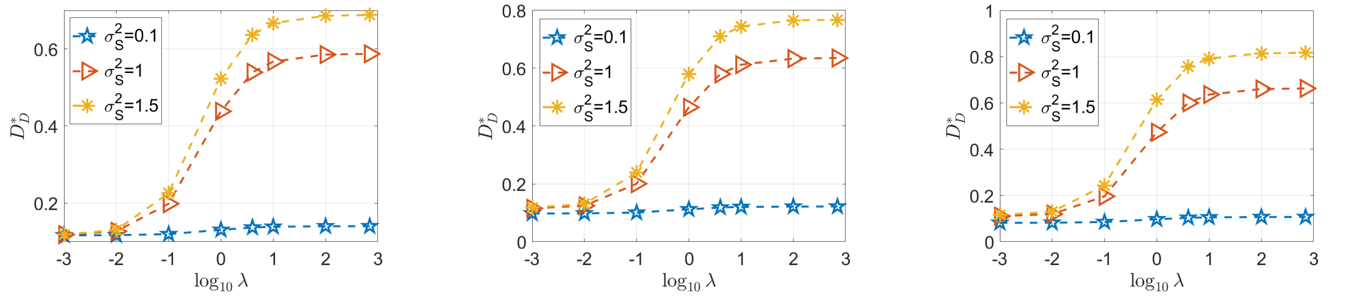
For a given correlation ρ, we observe in Fig. 3 that as σ_S^2 decreases, the receiver objective decreases. As the variance of the bias σ_S^2 decreases, the sender's opinions are closer to their true value, and the objectives of the sender and the receiver become more aligned.

When the bias is negligible ($\sigma_S \to 0$), the objectives of all the senders are similar, resulting in a negligible change in the receiver objective with λ, which we observe in Fig. 3 for $\sigma_S^2 = 0.1$.

4.3 Comparison with different types of senders

In Fig. 4, the receiver objective for the following four different types of senders is plotted for a specific setting with $\sigma_X^2 = \sigma_S^2 = 1, \rho = 0.5$:

(1) non-strategic (S_n): All agents are non-strategic and send their honest reply (E_0).
(2) full information (S_f): All agents are fully rational and have full information. The classifier here is that in Anand and Akyol (2024).
(3) bounded rational (S_b): The agents follow the setting described in this paper.
(4) partially-strategic (S_p): All agents minimize $\mathbb{E}\{(X+S-Y)^2\}$, but they assume the receiver is not strategic and hence implements a naive estimator, $\mathbf{y}^{(0)}$. The classifier is the same as that for E_1.



(a) $\rho = 0.1$ (b) $\rho = 0.5$ (c) $\rho = 0.7$

Fig. 3. Receiver objective D_D^* for $M = 4$ classification of $(X, S) \sim \mathcal{N}\left(\begin{bmatrix}0\\0\end{bmatrix}, \begin{bmatrix}1 & \sigma_S\rho\\ \sigma_S\rho & \sigma_S^2\end{bmatrix}\right)$ for a given correlation ρ value with respect to bias variance σ_S^2.

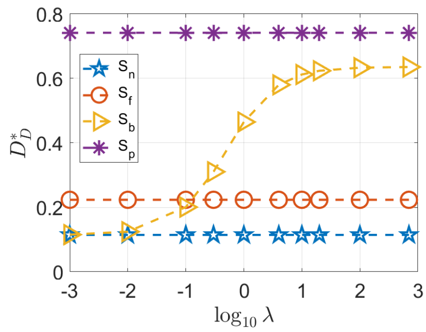


Fig. 4. Receiver objective D_D^* for $M = 4$ classification of $(X, S) \sim \mathcal{N}\left(\begin{bmatrix}0\\0\end{bmatrix}, \begin{bmatrix}1 & 0.5\\ 0.5 & 1\end{bmatrix}\right)$ for a full information fully rational estimator with four different types of senders.

The receiver is fully rational with full information about the type of sender, the source distribution, and sender and receiver objectives.

As expected, the non-strategic sender results in the lowest receiver objective. For negligible λ, the population is mostly of level-0 cognitive type, as mentioned before. Since they respond honestly, S_b is closer to the non-strategic value as $\lambda \to 0$.

Although we expect that S_f results in maximizing the receiver objective among the above four senders, we observe from Fig. 4 that the receiver may prefer a fully rational sender with full information to other types of strategic senders for $\lambda > \lambda_0$ for some λ_0. We observe that the receiver benefits from a boundedly or fully rational setup compared to the setting where all senders are partially-strategic.

5 Conclusions

In this paper, we analyzed the problem of strategic classification of a 2-dimensional source (X, S) with three types of senders with hierarchical cognitive types: level-0 non-strategic, level-1 strategic, and level-2 strategic, where each level assumes they are unique and that the other agents have lower cognitive levels. We considered quadratic objectives for the sender and the receiver and extended our prior work on design, a gradient-descent based algorithm for a 2-dimensional source with a single type of fully rational sender with full information, to the bounded-rational setting considered in this paper. The numerical results obtained via the proposed algorithm suggest several intriguing research problems that we leave as a part of our future work.

References

Akyol, E. and Anand, A. (2023). Strategic Quantization. In *IEEE International Symposium on Information Theory (ISIT)*, 543–548.

Anand, A. and Akyol, E. (2024). Strategic Quantization with Quadratic Distortion Measures. In *https://tinyurl.com/CPHS2024-bounded-rationality, Working Paper*.

Bénabou, R. and Tirole, J. (2006). Incentives and Prosocial Behavior. *American Economic Review*, 96(5), 1652–1678.

Camerer, C.F., Ho, T.H., and Chong, J.K. (2004). A Cognitive Hierarchy Model of Games. *The Quarterly Journal of Economics*, 119(3), 861–898.

Crawford, V.P. and Sobel, J. (1982). Strategic Information Transmission. *Econometrica: Journal of the Econometric Society*, 1431–1451.

Fischer, P.E. and Verrecchia, R.E. (2000). Reporting Bias. *The Accounting Review*, 75(2), 229–245.

Kamenica, E. and Gentzkow, M. (2011). Bayesian Persuasion. *American Economic Review*, 101(6), 2590–2615.

Rayo, L. and Segal, I. (2010). Optimal Information Disclosure. *Journal of Political Economy*, 118(5), 949–987.