Signaling games in multiple dimensions: Geometric properties of equilibrium solutions☆

Ertan Kazıklı [a],*, Sinan Gezici [b], Serdar Yüksel [c]

[a] *Electrical and Electronics Engineering, TOBB University of Economics and Technology, Ankara, Turkey*
[b] *Electrical and Electronics Engineering, Bilkent University, Ankara, Turkey*
[c] *Mathematics and Statistics, Queen's University, Kingston, Ontario, Canada*

ARTICLE INFO

ABSTRACT

Signaling game problems investigate communication scenarios where encoder(s) and decoder(s) have misaligned objectives due to the fact that they either employ different cost functions or have inconsistent priors. This problem has been studied in the literature for scalar sources under various setups. In this paper, we consider multi-dimensional sources under quadratic criteria in the presence of a bias leading to a mismatch in the criteria, where we show that the generalization from the scalar setup is more than technical. We show that the Nash equilibrium solutions lead to structural richness due to the subtle geometric analysis the problem entails, with consequences in both system design, the presence of linear Nash equilibria, and an information theoretic problem formulation. We first provide a set of geometric conditions that must be satisfied in equilibrium considering any multi-dimensional source. Then, we consider independent and identically distributed sources and characterize necessary and sufficient conditions under which an informative linear Nash equilibrium exists. These conditions involve the bias vector that leads to misaligned costs. Depending on certain conditions related to the bias vector, the existence of linear Nash equilibria requires sources with a Gaussian or a symmetric density. Moreover, in the case of Gaussian sources, our results have a rate–distortion theoretic implication that achievable rates and distortions in the considered game theoretic setup can be obtained from its team theoretic counterpart.

1. Introduction

In a team theoretic setup where the decision makers share a common goal, the decision makers do not wish to hide information to improve the performance since revealing more information does not lead to a degradation of system performance. Therefore, in such setups, if there is no constraint on messages to transmit between the decision makers, such as a power constraint or a limited bandwidth requirement, a decision maker can always reveal more information without causing any performance loss. On the other hand, in a game theoretic (strategic) setup involving decision makers with misaligned goals, revealing more information may hurt some or even all of the decision makers (Bassan, Gossner, Scarsini, & Zamir, 2003). Hence, a decision maker in

a strategic setting needs to take misaligned goals into account while designing what information to reveal to another decision maker. We may consider two main themes which lead to misaligned objectives for the decision makers. In the first theme, the decision makers employ different cost functions, e.g., a decision maker wishes to mislead another decision maker, see, e.g., Akyol, Langbort, and Başar (2017), Sarıtaş, Yüksel, and Gezici (2017) and Le Treust and Tomala (2019). The second theme is concerned with the case when the decision makers have subjective beliefs regarding prior probability distributions of unknown parameters. This subjectivity leads to misaligned objectives for the decision makers even though they employ the same cost function, see, e.g., Başar (1985), Kazıklı, Sarıtaş, Gezici, and Yüksel (2023) and Sarıtaş, Gezici, and Yüksel (2019). These both lead to a game theoretic setup where a suitable equilibrium concept, such as the Nash equilibrium and the Stackelberg equilibrium, is to be used to analyze the system. These problems fall into the general class of signaling game problems that investigates communication scenarios between decision makers with misaligned objectives. In this context, Crawford and Sobel, in their seminal paper (Crawford & Sobel, 1982), introduce a signaling game problem where a biased encoder wishes to convey a scalar source to a decoder, and

a message transmission does not induce a cost for the encoder. This problem is also referred to as *cheap talk*, which emphasizes that communication is costless. Crawford and Sobel show that under certain technical conditions regarding cost functions, the encoder must hide information at a Nash equilibrium by employing quantization policies, which holds even though there is no restriction on communication. In an equilibrium with a quantization policy, referred to as quantized or partition equilibrium, the encoder partitions the observation space into intervals and reveals the interval that contains the encoder's observation. Crawford and Sobel's result implies that at a Nash equilibrium, the encoder cannot convey its private information completely by employing a linear encoding policy (i.e., transmitting a scaled version of its observation to the decoder). This is a striking example where providing more information to the decoder by employing a linear encoder instead of a quantized encoder breaks the equilibrium in a game theoretic setup. In this manuscript, we study multi-dimensional sources under quadratic criteria for the cheap talk setup of Crawford and Sobel and investigate the properties of Nash equilibrium solutions.

Our work investigates communication scenarios between a biased encoder and a decoder, which leads to a signaling game problem. We may encounter biased decision makers in various applications. For instance, in control applications, an adversary may wish to inject a bias into a control system in order to deteriorate the system performance (Teixeira, Shames, Sandberg, & Johansson, 2015). In smart grid applications, a strategic consumer or electricity producer in a microgrid system may wish to give false or biased measurement reports to another decision maker for its own benefit (Larrousse, Beaude, & Lasaulce, 2014). As another application, strategic users in a cellular network may wish to misreport their channel conditions to the base station for their own benefit (Kavitha, Altman, El-Azouzi, & Sundaresan, 2012). Moreover, interactions between attackers and defenders in control applications may be modeled as a cheap talk problem (Li, Dán, & Liu, 2020; Sarıtaş, Dán and Sandberg, 2020). For applications of signaling games and cheap talk in fields such as economics, finance, biology, and political science, the reader is referred to Sobel (2020).

1.1. Preliminaries

We consider the following multi-dimensional signaling game problem where an encoder and a decoder communicate. This, in particular, corresponds to a multi-dimensional cheap talk problem where cheap talk refers to the fact that a message transmission does not induce a cost for the encoder. The encoder observes the value of an n-dimensional random vector $\boldsymbol{M} = [M_1, \ldots, M_n]^T$ where $M_1, \ldots, M_n$ are $\mathbb{M}$-valued random variables. The encoder conveys a message $\boldsymbol{Z} = [Z_1, \ldots, Z_n]^T$ via an encoding policy $\gamma^e(\cdot)$, i.e., $\boldsymbol{Z} = \gamma^e(\boldsymbol{M})$, where $Z_1, \ldots, Z_n$ are $\mathbb{Z}$-valued random variables. The decoder directly observes $\boldsymbol{Z}$ and takes an action $\boldsymbol{U} = [U_1, \ldots, U_n]^T$ via a decoding policy $\gamma^d(\cdot)$, i.e., $\boldsymbol{U} = \gamma^d(\boldsymbol{Z})$, where $U_1, \ldots, U_n$ are $\mathbb{M}$-valued random variables. In this paper, we consider real valued random variables, i.e., $\mathbb{M} = \mathbb{Z} = \mathbb{R}$ where $\mathbb{R}$ denotes the set of real numbers. The aim of the encoder is to minimize $J^e(\gamma^e, \gamma^d) = \mathbb{E}[c^e(\boldsymbol{M}, \boldsymbol{U})]$ where[1]

$$c^e(\boldsymbol{m}, \boldsymbol{u}) = \sum_{i=1}^{n}(m_i - u_i - b_i)^2 = \|\boldsymbol{m} - \boldsymbol{u} - \boldsymbol{b}\|^2. \tag{1}$$

In (1), $\boldsymbol{b}$ denotes a deterministic bias vector which is common knowledge among the players and quantifies the degree of misalignment between the objective functions of the encoder and

decoder. In other words, the encoder wishes to make biased reports regarding its observations possibly with different biases for different components. On the other hand, the decoder wishes to estimate the random source vector as accurately as possible; thus, its objective function does not include a bias vector. In particular, the aim of the decoder is to minimize $J^d(\gamma^e, \gamma^d) = \mathbb{E}[c^d(\boldsymbol{M}, \boldsymbol{U})]$ where

$$c^d(\boldsymbol{m}, \boldsymbol{u}) = \sum_{i=1}^{n}(m_i - u_i)^2 = \|\boldsymbol{m} - \boldsymbol{u}\|^2. \tag{2}$$

The communication scenario is depicted in Fig. 1. Our aim is to characterize the Nash equilibrium where the decision makers announce their policies at the same time. At a Nash equilibrium, none of the players wishes to unilaterally deviate from their current strategies as their cost cannot get better by doing so. In particular, a set of policies $\gamma^{*,e}$ and $\gamma^{*,d}$ forms a Nash equilibrium (e.g., Başar & Olsder, 1999) if

$$J^e(\gamma^{*,e}, \gamma^{*,d}) \le J^e(\gamma^e, \gamma^{*,d}) \text{ for all } \gamma^e \in \Gamma^e,$$
$$J^d(\gamma^{*,e}, \gamma^{*,d}) \le J^d(\gamma^{*,e}, \gamma^d) \text{ for all } \gamma^d \in \Gamma^d, \tag{3}$$

where Γ^e and Γ^d are the sets of all deterministic (and Borel measurable) functions from $\mathbb{M}^n$ to $\mathbb{Z}^n$ and from $\mathbb{Z}^n$ to $\mathbb{M}^n$, respectively.

Remark 1. Under the Nash equilibrium concept, both players announce their policies at the same time. By considering the Nash equilibrium concept, we essentially investigate a non-cooperative communication setup in terms of policy announcements in the sense that no player discloses its policy before the other player. This means that no player commits to a certain announced policy *a priori*. This equilibrium concept is appropriate, for instance, when the players do not have access to policy announcements of each other or when they do not trust an announced policy by the other player. In contrast, one can also consider the Stackelberg setup (see, e.g., Başar and Olsder (1999) for a definition) where the encoder announces its policy and commits to this policy, and the decoder chooses its policy given the encoder's announcement. We may view the Stackelberg setup as a cooperative communication setup as there is a policy announcement by the encoder. In fact, for the scalar or multi-dimensional cheap talk setup, the Stackelberg equilibrium solution leads to full revelation where the encoder discloses the source completely (Sarıtaş et al., 2017, Theorem 3.3). In contrast, there does not exist a Nash equilibrium with full information revelation in general. In other words, the encoder must hide information partially (or even completely in certain cases, see, e.g., Kazıklı, Sarıtaş, Gezici, Linder, and Yüksel (2022, Theorem 3) for the scalar case) in the non-cooperative communication setup whereas it does not hide any information in the cooperative communication setup.

Assumption 2. Considering each component M_i of the source random vector $\boldsymbol{M}$, every non-empty open set on its support has a positive measure.

The following is an implication of this assumption. Consider a convex set C with a non-empty interior. Then, its centroid $\mathbb{E}[\boldsymbol{M}|\boldsymbol{M} \in C]$ must be in the interior of set C.[2] We will use this implication later in the paper.

We formally define a quantization policy in the following. Note that due to results in Crawford and Sobel (1982), Sarıtaş et al. (2017), a Nash equilibrium in the scalar source case must involve quantization policies at the encoder with convex bins.

---

[1] We adopt the convention that random variables are denoted by uppercase letters with their realizations denoted by the corresponding lowercase letters.

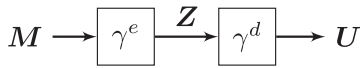[2] This follows from a separating hyperplane argument.

Fig. 1. Communication setting.

Definition 3. A quantization policy with K bins, q, is a (Borel) measurable mapping from $\mathbb{M}^n = \mathbb{R}^n$ to the set $\{1, \ldots, K\}$ characterized by a measurable partition $\{\mathcal{B}^1, \ldots, \mathcal{B}^K\}$ such that $\mathcal{B}^i = \{\boldsymbol{m} \mid q(\boldsymbol{m}) = i\}$ for $i = 1, \ldots, K$ and that bin probabilities are strictly positive. The $\mathcal{B}^i$ are called the bins of q.

The bins defined in Definition 3 lead to a Nash equilibrium under certain conditions described later in the manuscript. If these bins form a Nash equilibrium, they are referred to as (Nash) equilibrium partitions. In contrast to the scalar source case, there may exist a Nash equilibrium with a linear encoder in the multi-dimensional source case, which is investigated later in the paper. Accordingly, we make the following definition.

Definition 4. For the n-dimensional cheap talk problem, if an encoding policy $\boldsymbol{z} = \gamma^e(\boldsymbol{m}) = A\boldsymbol{m}$ where $A \in \mathbb{R}^{m \times n}$ with $m \leq n$ and a decoding policy $\boldsymbol{u} = \gamma^d(\boldsymbol{z})$ satisfy (3), we say that these policies lead to a linear Nash equilibrium.

Definition 5. We say that a Nash equilibrium is *informative* if the encoder reveals information related to the source, i.e., the source $\boldsymbol{M}$ and the message $\boldsymbol{Z}$ are not independent random variables. A Nash equilibrium is referred to as *non-informative* when the encoded message is independent of the source.

We note that there always exists a non-informative Nash equilibrium for the multi-dimensional cheap talk problem, which follows from Crawford and Sobel (1982). In this equilibrium, the encoder transmits a message which is independent of the source. The decoder takes an action based on the prior probability distribution of the source, i.e., its best response $\boldsymbol{u} = \mathbb{E}[\boldsymbol{M}]$. This is a Nash equilibrium since both the encoder and the decoder cannot improve their expected costs by deviating from these strategies. In contrast, an informative Nash equilibrium may or may not exist depending on the setup.

At a given Nash equilibrium, all possible realized values of $\boldsymbol{u}$ are referred to as *decoder actions*. While investigating the geometric properties of Nash equilibria, we frequently use the following definition regarding the set of decoder actions in equilibrium.

Definition 6. We say that a non-empty set of decoder actions containing more than one element forms a *continuum* if it is a closed and connected set (i.e., it cannot be expressed as a union of two or more disjoint and closed sets).

An important implication of our results is related to the information theoretic limits of the cheap talk problem. In classical communication settings involving decision makers with aligned goals, information theoretic limits specify bounds on the rate of communication and system performance measured by a common cost criterion (see, e.g., Cover and Thomas (2006)). In such settings, a bound on the achievable communication rate arises due to system requirements such as a power constraint at the encoder and having a noisy channel. On the other hand, an interesting question arises in a game theoretic setup: Does an upper bound exist on the achievable rate of communication due to misaligned cost criteria? In certain cases, our analysis gives a conclusive answer to this question for the multi-dimensional cheap talk setup. In particular, we show that there exists a Nash equilibrium with a linear encoder depending on certain explicit conditions, in which case there does not exist an upper bound on the achievable rate of communication. We consider the Nash setup for such an information theoretic problem. We refer the reader to Le Treust and Tomala (2019) for a Stackelberg (Bayesian persuasion) setup.

1.2. Literature review

The cheap talk and signaling game problems have gained significant attention in recent control and communication theory literature. For instance, Sarıtaş et al. (2017) investigate signaling game setups with quadratic cost criteria under Nash and Stackelberg equilibria concepts where a biased encoder communicates with a decoder. The work in Akyol et al. (2017) considers a Gaussian signaling game problem under the Stackelberg equilibrium concept where the bias term at the encoder is modeled as a random variable. In Sayın, Akyol, and Başar (2019), a multi-stage Gaussian signaling setup is investigated under the Stackelberg equilibrium concept where the private state of the encoder is a controlled Gauss–Markov process. The work in Le Treust and Tomala (2019) investigates information theoretic limits for the Bayesian persuasion (Stackelberg) setup where there is a commitment assumption for the encoder. The works in Vora and Kulkarni (2020a, 2020b) consider problems under the Stackelberg equilibrium concept where the decoder has a commitment assumption and introduce the notion of information extraction capacity. In Kazıklı et al. (2022), various properties of Nash equilibria are analyzed for the one-dimensional quadratic cheap talk problem. In Sarıtaş, Yüksel and Gezici (2020), multi-stage cheap talk and signaling game problems are investigated under Nash and Stackelberg equilibria. In Kazıklı, Gezici, and Yüksel (2021), some of the preliminary results in this paper were announced, and the results presented did not include proofs.

Multi-dimensional cheap talk problems have also been considered in the economics literature (Ambrus & Takahashi, 2008; Battaglini, 2002; Chakraborty & Harbaugh, 2007; Levy & Razin, 2007; Miura, 2014). For instance, Levy and Razin (2007) investigate a two-dimensional source setting where an encoder communicates with a decoder. Different from our work, the encoder's preferences over different decoder actions are primarily determined by preferences in a certain dimension. In particular, if the encoder prefers one decoder action over the other in this dimension, then the second dimension does not matter. In this case, Levy and Razin (2007) show the existence of an upper bound on the number of decoder actions. In addition, the work in Battaglini (2002) considers a multi-dimensional cheap talk problem with two encoders and a decoder. While Battaglini (2002) studies conditions on the existence of equilibria with the encoders completely revealing their observations, our focus instead is on the characterization of Nash equilibrium partitions in general; as in the case with a single encoder, we do not have full revelation in general. More specifically, we focus on a scenario with a single encoder that jointly encodes its multi-dimensional observation and employs a single quadratic cost function. More recently, Sémirat (2019) investigates a two-dimensional cheap talk setup between an encoder and a decoder considering a uniform source where the encoder is restricted to transmit a binary message. This work proves the existence of an informative Nash equilibrium for any bias vector under the considered setup.

1.3. Contributions

The main aim of this paper is to analyze a quadratic multi-dimensional cheap talk problem, which is a multi-dimensional extension of Crawford and Sobel's formulation (Crawford & Sobel, 1982). The main contributions of this paper can be summarized as follows:

(i) We show that for general source distributions, decoder actions in any Nash equilibrium must satisfy a necessary geometric condition (Lemma 7).

(ii) We derive the necessary conditions that a Nash equilibrium with a continuum of decoder actions needs to satisfy in the case of two-dimensional observations with general distributions (Lemmas 11 and 12).

(iii) We completely characterize necessary and sufficient conditions under which linear Nash equilibria exist considering independent and identically distributed (i.i.d.) two-dimensional observations (Theorem 13). We also generalize these results to the case when the encoder makes more than two i.i.d. observations (Theorems 17 and 19).

(iv) We take the dimension of the source process to infinity and provide an information theoretic perspective to the cheap talk problem by introducing a rate–distortion theoretic formulation. We obtain achievable rates and distortions for the particular case of i.i.d. Gaussian sources (Theorem 21).

2. Geometric properties of Nash equilibria

2.1. A necessary geometric condition for Nash equilibria

In this subsection, we show that the cost structure employed in the problem imposes certain restrictions on the actions taken by the decoder at a Nash equilibrium. In particular, we derive a geometric condition that any two decoder actions at a Nash equilibrium must satisfy. This derivation also allows us to specify the general structure of a Nash equilibrium with a quantization policy at the encoder. In addition, while this geometric condition is important on its own as it provides a necessary condition for a Nash equilibrium in terms of induced decoder actions, it is also useful while deriving conditions for the existence of linear Nash equilibria. It is noted that the following result holds regardless of the source distribution and applies to both i.i.d. and non-i.i.d. sources.

Lemma 7. *Consider the n-dimensional cheap talk problem with the source random vector $\boldsymbol{M} = [M_1, \ldots, M_n]^T$ where each element of $\boldsymbol{M}$ can have different distributions and can be dependent or independent. Let $\mathcal{B}^\alpha$ and $\mathcal{B}^\beta$ be two bins, and let $\boldsymbol{u}^\alpha = \mathbb{E}[\boldsymbol{M}|\boldsymbol{M} \in \mathcal{B}^\alpha]$ and $\boldsymbol{u}^\beta = \mathbb{E}[\boldsymbol{M}|\boldsymbol{M} \in \mathcal{B}^\beta]$ denote their centroids which are the decoder actions taken when the encoder reveals $\boldsymbol{M} \in \mathcal{B}^\alpha$ and $\boldsymbol{M} \in \mathcal{B}^\beta$, respectively.*

(i) *These decoder actions must satisfy the following necessary condition at a Nash equilibrium:*

$$2 \left| (\boldsymbol{u}^\beta - \boldsymbol{u}^\alpha)^T \boldsymbol{b} \right| \leq \| \boldsymbol{u}^\beta - \boldsymbol{u}^\alpha \|^2. \tag{4}$$

(ii) *At a Nash equilibrium, the encoder decomposes the complete observation space into two regions via a hyperplane orthogonal to $(\boldsymbol{u}^\alpha - \boldsymbol{u}^\beta)$ and intersecting the line connecting $\boldsymbol{u}^\alpha$ and $\boldsymbol{u}^\beta$, and $\mathcal{B}^\alpha$ and $\mathcal{B}^\beta$ are subsets of these respective regions. In particular, $\mathcal{B}^\alpha$ must be a subset of the set $\{\boldsymbol{m} \mid h(\boldsymbol{m}, \boldsymbol{u}^\alpha, \boldsymbol{u}^\beta) \geq 0\}$ whereas $\mathcal{B}^\beta$ must be a subset of the set $\{\boldsymbol{m} \mid h(\boldsymbol{m}, \boldsymbol{u}^\alpha, \boldsymbol{u}^\beta) \leq 0\}$ where*

$$h(\boldsymbol{m}, \boldsymbol{u}^\alpha, \boldsymbol{u}^\beta) \triangleq$$
$$\left(\boldsymbol{m} - \left(\frac{\boldsymbol{u}^\beta + \boldsymbol{u}^\alpha}{2} + \boldsymbol{b} \right) \right)^T (\boldsymbol{u}^\beta - \boldsymbol{u}^\alpha), \tag{5}$$

and $h(\boldsymbol{m}, \boldsymbol{u}^\alpha, \boldsymbol{u}^\beta) = 0$ defines the hyperplane on which the encoder is indifferent between either decoder actions, i.e., these $\boldsymbol{m}$ values may belong to both $\mathcal{B}^\alpha$ and $\mathcal{B}^\beta$.

(iii) *At a Nash equilibrium where the encoder uses quantization policies, the quantization bins are always convex.*

See Appendix A for a proof. Fig. 2 illustrates the result in Lemma 7 for an example setup. In the case of more than two decoder actions, each pair of decoder actions must satisfy the

condition in (4) at a Nash equilibrium. In addition, the bins for each decoder action must be obtained by computing half spaces via (5) for each pair of decoder actions and then by intersecting these half spaces. In particular, if the decoder actions $\{\boldsymbol{u}^1, \ldots, \boldsymbol{u}^K\}$ and the corresponding bins $\{\mathcal{B}^1, \ldots, \mathcal{B}^K\}$ form a Nash equilibrium with K bins, then it must be that

$$\mathcal{B}^i = \{\boldsymbol{m} \mid h(\boldsymbol{m}, \boldsymbol{u}^i, \boldsymbol{u}^j) \geq 0 \text{ for all } j \neq i\}, \tag{6}$$

for $i = 1, \ldots, K$. Note that the conditions in (6) are necessary but not sufficient for a Nash equilibrium. Due to the equilibrium conditions at the decoder, for a Nash equilibrium with K bins, the decoder actions $\{\boldsymbol{u}^1, \ldots, \boldsymbol{u}^K\}$ and the corresponding bins $\{\mathcal{B}^1, \ldots, \mathcal{B}^K\}$ must also satisfy the following centroid conditions:

$$\boldsymbol{u}^i = \mathbb{E}[\boldsymbol{M}|\boldsymbol{M} \in \mathcal{B}^i], \tag{7}$$

for $i = 1, \ldots, K$. If the conditions in (6) and (7) are satisfied, then the corresponding decoder actions and bins form a Nash equilibrium with K bins. Fig. 3 depicts a Nash equilibrium involving a quantization policy with three bins at the encoder.

Remark 8. In the case of scalar cheap talk, it is required that $|u^\alpha - u^\beta| > 2|b|$ holds for any decoder actions u^α and u^β at a Nash equilibrium. This directly implies that a Nash equilibrium must involve quantization policies as concluded in Sarıtaş et al. (2017, Theorem 3.2). In contrast, such a direct conclusion does not hold for the multi-dimensional cheap talk problem. In fact, if $(\boldsymbol{u}^\beta - \boldsymbol{u}^\alpha)$ is orthogonal to $\boldsymbol{b}$, then the necessary condition in (4) is always satisfied regardless of the distance between these two decoder actions. This permits the existence of linear Nash equilibria when the source is multi-dimensional, depending on certain conditions investigated later in the paper. Since such an orthogonality property does not hold when the source is one-dimensional, decoder actions cannot get arbitrarily close. Hence, there does not exist a linear Nash equilibrium in this case.

Lemma 7 presents a geometric condition that any two decoder actions at a Nash equilibrium must satisfy. It is important to emphasize that this condition applies to any joint distribution for multi-dimensional observations. In particular, Lemma 7 holds even for joint distributions that are not independent and identically distributed.

2.2. Necessary conditions for continuum of decoder actions in equilibrium

In this subsection, we further investigate the geometric condition in Lemma 7 to derive conditions that a Nash equilibrium with a connected set of decoder actions must satisfy for the particular case of two-dimensional cheap talk. Since a linear encoding policy induces a connected set of decoder actions, our results in this subsection are useful while deriving conditions for the existence of a linear Nash equilibrium.

Lemma 7 implies that for decoder actions $\boldsymbol{u}^\alpha$ and $\boldsymbol{u}^\beta$ satisfying $(\boldsymbol{u}^\beta - \boldsymbol{u}^\alpha)^T \boldsymbol{b} = 0$, it is possible to make their distance $\| \boldsymbol{u}^\alpha - \boldsymbol{u}^\beta \|$ arbitrarily small. On the other hand, for decoder actions $\boldsymbol{u}^\alpha$ and $\boldsymbol{u}^\beta$ with $(\boldsymbol{u}^\beta - \boldsymbol{u}^\alpha)^T \boldsymbol{b} \neq 0$, since the distance $\| \boldsymbol{u}^\alpha - \boldsymbol{u}^\beta \|$ is lower bounded by a positive value, these decoder actions $\boldsymbol{u}^\alpha$ and $\boldsymbol{u}^\beta$ cannot get arbitrarily close. This motivates an equivalent formulation by introducing the following transformation of variables. In particular, we define

$$\boldsymbol{X} = \mathcal{T}\boldsymbol{M}, \tag{8}$$

$$\boldsymbol{U} = \mathcal{T}^{-1}\boldsymbol{Y}, \tag{9}$$

where

$$\mathcal{T} = \begin{bmatrix} -b_2 & b_1 \\ b_1 & b_2 \end{bmatrix}, \quad \mathcal{T}^{-1} = \frac{1}{b_1^2 + b_2^2} \begin{bmatrix} -b_2 & b_1 \\ b_1 & b_2 \end{bmatrix}, \tag{10}$$
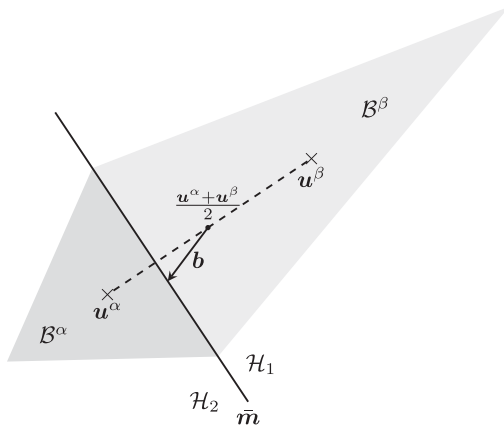
Fig. 2. Illustration of half spaces induced by decoder actions $\boldsymbol{u}^\alpha$ and $\boldsymbol{u}^\beta$ in Lemma 7. The crosses represent the decoder actions, and the arrow represents the bias vector. These decoder actions and the bias vector lead to a line of $\bar{\boldsymbol{m}}$ values for which the encoder is indifferent between reporting these observations as $\boldsymbol{u}^\alpha$ and $\boldsymbol{u}^\beta$. The shaded areas illustrate example bins which satisfy the necessary condition that the half spaces $\mathcal{H}_1$ and $\mathcal{H}_2$ cannot intersect with $\mathcal{B}^\alpha$ and $\mathcal{B}^\beta$, respectively.
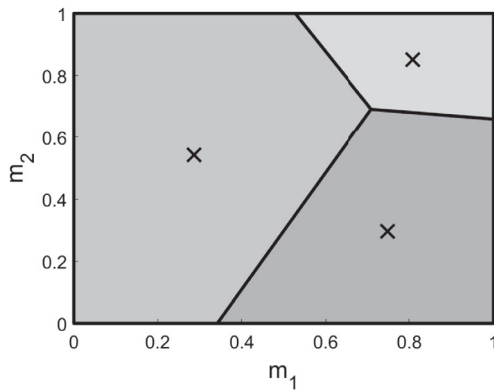

Fig. 3. Illustration of a Nash equilibrium involving a quantization policy with three bins for the case when $b_1 = b_2 = 0.1$ and the source is two-dimensional i.i.d. with a uniform distribution. The shaded areas show the quantization bins, the lines between the areas are the bin edges, and the crosses represent the decoder actions induced in equilibrium. These quantization bins and bin edges satisfy (6) and (7), which leads to a Nash equilibrium.
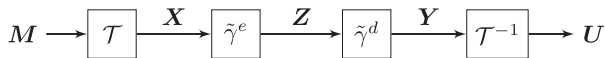

Fig. 4. Equivalent formulation where $\mathcal{T}$ denotes the linear transformation specified by $X_1 = b_1 M_2 - b_2 M_1$ and $X_2 = b_1 M_1 + b_2 M_2$, and $\mathcal{T}^{-1}$ denotes its inverse.

and $\boldsymbol{X} \triangleq [X_1, X_2]^T$ and $\boldsymbol{Y} \triangleq [Y_1, Y_2]^T$ respectively denote the observation at the encoder and the decoder action in the transformed coordinate system. The proposed equivalent formulation is depicted in Fig. 4 where the linear transformation $\mathcal{T}$ and its inverse $\mathcal{T}^{-1}$ are fixed, and the encoder and decoder design $\tilde{\gamma}^e(\cdot)$ and $\tilde{\gamma}^d(\cdot)$, respectively. In the following lemma, we show that the proposed transformation of variables leads to an equivalent formulation. See Appendix B for a proof.

Lemma 9. *Suppose that the encoder uses a fixed transformation from the source $\boldsymbol{M}$ to an auxiliary variable $\boldsymbol{X}$ specified by (8) and designs the map $\tilde{\gamma}^e(\cdot)$ from $\boldsymbol{X}$ to the encoded message $\boldsymbol{Z}$. Suppose that the decoder designs the map $\tilde{\gamma}^d(\cdot)$ from its observation $\boldsymbol{Z}$ to an auxiliary variable $\boldsymbol{Y}$ and employs a fixed transformation from $\boldsymbol{Y}$*

to the decoder action $\boldsymbol{U}$ specified by (9). Then, designing $\tilde{\gamma}^e(\cdot)$ at the encoder and $\tilde{\gamma}^d(\cdot)$ at the decoder is equivalent to the original problem where the encoder designs the map $\gamma^e(\cdot)$ from $\boldsymbol{M}$ to $\boldsymbol{Z}$ under the cost criterion (1), and the decoder designs the map $\gamma^d(\cdot)$ from $\boldsymbol{Z}$ to $\boldsymbol{U}$ under the cost criterion (2). In particular, an equilibrium under the proposed formulation is also an equilibrium under the problem given in (3) and vice versa. In this equivalent formulation, the aim of the encoder and decoder is to minimize $\tilde{J}^e(\tilde{\gamma}^e, \tilde{\gamma}^d) \triangleq \mathbb{E}[c_t^e(\boldsymbol{X}, \boldsymbol{Y})]$ and $\tilde{J}^d(\tilde{\gamma}^e, \tilde{\gamma}^d) \triangleq \mathbb{E}[c_t^d(\boldsymbol{X}, \boldsymbol{Y})]$, respectively, where

$$c_t^e(\boldsymbol{x}, \boldsymbol{y}) \triangleq (x_1 - y_1)^2 + (x_2 - y_2 - \tilde{b})^2 = c^e(\boldsymbol{m}, \boldsymbol{u})\tilde{b}, \tag{11}$$

$$c_t^d(\boldsymbol{x}, \boldsymbol{y}) \triangleq (x_1 - y_1)^2 + (x_2 - y_2)^2 = c^d(\boldsymbol{m}, \boldsymbol{u})\tilde{b}, \tag{12}$$

$$\tilde{b} \triangleq b_1^2 + b_2^2. \tag{13}$$

Lemma 10. *For a fixed encoding policy $\tilde{\gamma}^e(\boldsymbol{x})$, the optimal $\tilde{\gamma}^d(\cdot)$ that minimizes $\tilde{J}^d(\tilde{\gamma}^e, \tilde{\gamma}^d)$ is given by $\mathbb{E}[\boldsymbol{X}|\boldsymbol{Z} = \boldsymbol{z}]$.*

See Appendix C for a proof. Equipped with this equivalent formulation, we are now ready to present our results on necessary conditions for any Nash equilibrium with a continuum of decoder actions.

Lemma 11. *Consider the two-dimensional cheap talk problem. Suppose that at a given Nash equilibrium, a set of decoder actions $\mathcal{C}$ forms a continuum. Then, for any $\boldsymbol{y}^\alpha \in \mathcal{C}$ and $\boldsymbol{y}^\beta \in \mathcal{C}$, it must be that $y_2^\alpha = y_2^\beta$.*

See Appendix D for a proof. Lemma 11 implies that a continuum of actions is allowed only in a specific direction that depends on the bias terms in the original coordinate system.

Lemma 12. *Consider the two-dimensional cheap talk problem. Suppose that at a given Nash equilibrium, a set of decoder actions with the same second coordinate forms a continuum, i.e., $y_2 = \kappa$ where κ is in the support of X_2. Then, it must be that there exist decoder actions for all values of $y_1 \in [x_1^L(\kappa), x_1^U(\kappa)]$ and $y_2 = \kappa$ where $x_1^L(\kappa)$ and $x_1^U(\kappa)$ denote lower and upper boundaries of the support of X_1 when $X_2 = \kappa$, i.e., these decoder actions must be connected.*

See Appendix E for a proof. Lemma 11 states that a continuum of decoder actions must have a constant y_2 coordinate, and Lemma 12 states that this continuum of decoder actions must be supported *for all* values of y_1 in the support of X_1 given that $X_2 = y_2$. This means that a continuum of decoder actions cannot have a discontinuity. This type of continuum of actions can be attained by revealing the value of X_1 completely. In certain scenarios depending on the distribution and the bias vector, revealing X_1 can be a Nash equilibrium, as investigated in the next section.

3. Linear Nash equilibria

In this section, we present our main results on the existence of linear Nash equilibria. Towards that goal, we employ Lemmas 11 and 12 together with an interesting result from the literature known as Kagan–Linnik–Rao Theorem (Kagan, Linnik, & Rao, 1973, Theorem 5.3.1). We first consider the two-dimensional case in the following theorem. See Appendix F for a proof.

Theorem 13. *Consider the multi-dimensional cheap talk problem with sources M_1 and M_2, which are i.i.d. with the corresponding bias terms b_1 and b_2.*

(i) For $b_1 = 0$ or $b_2 = 0$, there always exists an informative Nash equilibrium with a linear encoder where the encoder completely reveals the source corresponding to a zero bias.

(ii) For $b_1 \neq 0$, $b_2 \neq 0$ and $|b_1| \neq |b_2|$, there exists an informative Nash equilibrium with a linear encoder if and only if the source distribution is Gaussian.

(iii) For $b_1 = b_2 \neq 0$, there exists an informative Nash equilibrium with a linear encoder if and only if the source distribution is symmetric around its mean, i.e., denoting the density of M_1 by $f(\cdot)$, we have that $f(\mu + x) = f(\mu - x)$ for almost all x where $\mathbb{E}[M_1] = \mu$.

(iv) For $b_1 = -b_2 \neq 0$, there always exists an informative Nash equilibrium with a linear encoder regardless of the source distribution.

In Fig. 5, we depict a linear Nash equilibrium for the case with a two-dimensional uniform source. This scenario corresponds to the third case in Theorem 13, where the source distribution is symmetric, and the biases are the same in each dimension.

Remark 14. Lemmas 11 and 12 require that at a Nash equilibrium, a continuum of decoder actions can only exist in the direction orthogonal to the bias vector $\boldsymbol{b}$ without any discontinuity considering the original coordinate system. If the source distribution is such that $\mathbb{E}[X_2|X_1 = x_1] = \mathbb{E}[b_1M_1 + b_2M_2|b_1M_2 - b_2M_1 = x_1] = 0$ holds for all x_1, then an encoding policy $z = \gamma^e(\boldsymbol{m}) = b_1m_2 - b_2m_1$ leads to a continuum of decoder actions that satisfies the necessary conditions in Lemmas 11 and 12. In addition, such an encoding policy leads to a Nash equilibrium, as the proof of Theorem 13 reveals.

Remark 15. Theorem 13 shows that depending on certain conditions, there exists an informative Nash equilibrium with a linear encoder even for large values of $|b_1|$ and $|b_2|$. On the other hand, in the case of one-dimensional cheap talk, there may exist an upper bound on the number of bins in equilibrium, e.g., for sources with a bounded support (Crawford & Sobel, 1982) or for log-concave sources with a semi-unbounded support depending on certain conditions (Kazıklı et al., 2022). In addition, if the bias term is large, this upper bound may even be equal to one, which means that there does not exist an informative Nash equilibrium. Hence, even though the only Nash equilibrium in the case of a one-dimensional scenario may be non-informative, in the case of a two-dimensional scenario with the same bias as in the one-dimensional scenario in both dimensions, it is possible to obtain an informative Nash equilibrium when the source distribution is i.i.d. symmetric.

Remark 16. In the case of a Gaussian source, the problem decouples into two one-dimensional cheap talk problems. In particular, $X_1 = b_1M_2 - b_2M_1$ and $X_2 = b_1M_1 + b_2M_2$ become independent random variables when M_1 and M_2 are i.i.d. Gaussian. In fact, due to Darmois–Skitovich Theorem (Kagan et al., 1973, Theorem 3.1.1), X_1 and X_2 are independent only when M_1 and M_2 are Gaussian. As a result, the problem reduces to obtaining Nash equilibria for decoupled two one-dimensional cheap talk problems where an encoder wishes to convey X_1 with a zero bias and another encoder wishes to convey X_2 with a bias of $\tilde{b}$. From Kazıklı et al. (2022, Theorem 4), we know that in the case of one-dimensional cheap talk with a Gaussian source, for any $N \geq 1$, there exists a (unique) Nash equilibrium with N bins. Thus, for a two-dimensional cheap talk problem with a Gaussian source, there exists a Nash equilibrium where the encoder reveals X_1 completely and applies a quantization policy to X_2 with an arbitrary number of bins.

We can also consider n-dimensional i.i.d. Gaussian sources. In this case, one can apply an orthogonal transformation of variables in a similar manner to the two-dimensional case where random variables in each dimension are independent. Under this transformation of variables, there remains a bias term only for a single random variable. Due to Kazıklı et al. (2022, Theorem 4) and the independence of the random variables in the transformed coordinate system, it follows that there exists a Nash equilibrium where the encoder applies a quantization policy to this remaining random variable with any number of bins.

Theorem 17. Consider the n-dimensional cheap talk problem with an i.i.d. Gaussian source. Then, there exists a Nash equilibrium with a linear encoding policy where the encoder reveals all or a subset of $(n-1)$ dimensions completely (and applies a signaling game policy for the remaining dimension with any number of bins).

See Appendix G for a proof. Theorem 13 reveals that for the case with an i.i.d. Gaussian source, there always exists a linear Nash equilibrium regardless of the value of the bias vector. One can also consider non-i.i.d. Gaussian sources. In this case, we show that there may exist a linear Nash equilibrium depending on the bias vector and the covariance matrix in the following theorem, whose proof is presented in Appendix H.

Theorem 18. Consider the multi-dimensional cheap talk problem with Gaussian sources M_1 and M_2, and the corresponding bias terms b_1 and b_2. Let σ_1^2 and σ_2^2 denote the variances of M_1 and M_2, respectively, and let ρ denote their covariance. Then, there exists a Nash equilibrium with a linear encoding policy if $b_1b_2(\sigma_2^2 - \sigma_1^2) + (b_1^2 - b_2^2)\rho = 0$ holds.

Theorem 13 reveals that for the case with an i.i.d. two-dimensional symmetric source, there always exists an informative linear Nash equilibrium. When $n > 2$, it is possible to apply the linear policy in Theorem 13 for pairs of random variables to obtain a linear Nash equilibrium as we can obtain decoupled two-dimensional cheap talk problems. For instance, if $n = 2k$ for some $k > 2$, then revealing all or a subset of the random variables $M_2 - M_1, M_4 - M_3, \ldots, M_{2k} - M_{2k-1}$ yields a Nash equilibrium. In this case, the encoder reveals at most $n/2$ dimensions. In the following theorem, we show that a joint encoding policy can be applied to obtain a Nash equilibrium where the encoder reveals $(n-1)$ dimensions. The proof appears in Appendix I.

Theorem 19. Consider the n-dimensional cheap talk problem involving an i.i.d. source with a symmetric distribution. Then, there exists a Nash equilibrium with a linear encoding policy where the encoder reveals all or a subset of $(n-1)$ dimensions completely in a transformed coordinate system.

4. Large dimensions and a rate-distortion theoretic formulation of cheap talk

We have analyzed the multi-dimensional cheap talk problem where the bias vector at the encoder can be arbitrary. The special case when the components of the bias vector are the same leads to an important problem from an information theoretic perspective. In this case, the problem is to convey an i.i.d. source with a certain bias, and the bias is the same for each source component. In other words, the encoder observes independent copies from a random source and wishes to introduce the same bias for each independent copy. In such a problem, one may wish to obtain information theoretic limits of the communication. In a sense, this problem is a game theoretic counterpart of rate–distortion theory that is studied in a classical communication theoretic setup. Our findings reveal that if the distribution is Gaussian, then there always exists a linear Nash equilibrium where the encoder completely reveals $(n-1)$ dimensions in a transformed coordinate system. For the remaining dimension, the encoder has to
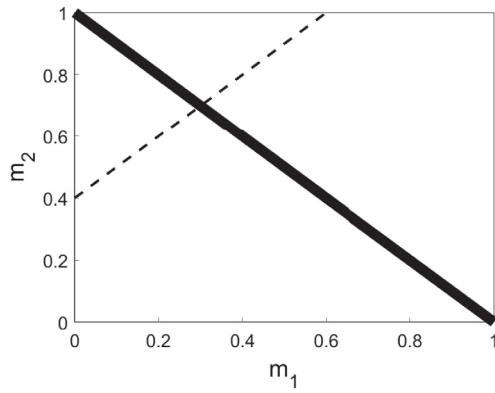
Fig. 5. Illustration of a Nash equilibrium with a linear encoder $\gamma^e(\boldsymbol{m}) = m_2 - m_1$ for the case when $b_1 = b_2 = 0.1$ and the source is two-dimensional i.i.d. with a uniform distribution where the solid line illustrates the continuum of decoder actions induced in equilibrium. The following interpretation can be made in relation to quantization policies (see also Fig. 3). When the encoder makes an observation exactly on the dashed line $m_2 - m_1 = 0.4$, the encoder only reveals that its observation is on this dashed line. The decoder takes the action $m_1 = 0.3$ and $m_2 = 0.7$ as its optimal response.

employ a signaling game policy with an arbitrary number of bins, including the case with one bin. This result holds because the problem can be transformed into decoupled problems consisting of a team theoretic problem for conveying an $(n-1)$-dimensional i.i.d. source without any bias and a one-dimensional cheap talk problem with a certain bias in the remaining dimension. If we increase the number of observed sources at the encoder, the effect of employed policy for this remaining dimension becomes negligible. This implies that the problem of finding achievable rate and distortion pairs is asymptotically equivalent to obtaining achievable rate and distortion pairs for a team theoretic setup in a transformed coordinate system.

The problem of interest is in fact can be more generally expressed in a rate–distortion theoretic formulation. The aim is to find the achievable rate and distortion region. In particular, we have the following problem:

Problem 20. *Consider the n-dimensional cheap talk problem with i.i.d. sources and $b = b_1 = \cdots = b_n$. We say that a tuple of rate and distortion pairs (R, D_e, D_d) is achievable at a Nash equilibrium if there exists a sequence of encoders and decoders that leads to a Nash equilibrium with the following properties:*

(i) The encoder is given by $\gamma_n^e : \mathbb{M}^n \to \{1, \dots, 2^{nR}\}$.
(ii) The decoder is given by $\gamma_n^d : \{1, \dots, 2^{nR}\} \to \mathbb{M}^n$ such that

$$\lim_{n \to \infty} \frac{\mathbb{E}\left[\sum_{i=1}^n (M_i - U_i - b)^2\right]}{n} \leq D_e, \tag{14}$$

$$\lim_{n \to \infty} \frac{\mathbb{E}\left[\sum_{i=1}^n (M_i - U_i)^2\right]}{n} \leq D_d. \tag{15}$$

Then, the problem is to determine if a given tuple (R, D_e, D_d) is achievable at a Nash equilibrium.

If the bias term is zero in this problem, then we obtain a team theoretic problem since the corresponding distortion values are identical at the encoder and decoder. We denote the corresponding rate and distortion values by R_T and D_T, respectively, where the subscript refers to the fact that the setup is team theoretic.

While we leave the study of Problem 20 for general sources for future work, the Gaussian case is completely solvable. Our result in the previous section shows that if the source distribution is Gaussian, one can apply a suitable transformation of variables to obtain an equivalent problem for which the encoder has a

bias only for a single random variable. We use this idea to relate achievable rate and distortion values of the original problem to that of a team theoretic problem. Before presenting this result, we note that at a Nash equilibrium we have $\mathbb{E}[\sum_{i=1}^n (M_i - U_i - b)^2] = \mathbb{E}[\sum_{i=1}^n (M_i - U_i)^2] + b^2 n$. Thus, we have the same rate region for any D_e value satisfying $D_e \geq D_d + b^2$. The following theorem characterizes achievable rates and distortion values for Problem 20 with Gaussian sources. See Appendix J for a proof.

Theorem 21. *Consider the multi-dimensional cheap talk problem with i.i.d. Gaussian sources where the bias term b is the same at each dimension. Suppose that a rate and a distortion pair (R_T, D_T) is achievable for the team theoretic problem with a zero bias. Then, for the game theoretic problem with a non-zero bias, the following rate and distortion values are achievable:*

$$R = R_T, \quad D_e \geq D_T + b^2, \quad D_d \geq D_T. \tag{16}$$

Remark 22. The proof of Theorem 17 reveals that by applying a joint encoding policy that uses multi-dimensional observations, it is possible to achieve team theoretic rates and distortions (except that there is still an additional b^2 term for the encoder's distortion). If we do not allow for joint encoding, which is equivalent to considering the scalar cheap talk setup, then the same rate in the team setup and in the game setup leads to different distortion values at the decoder, with the latter being larger.

In rate–distortion theory, an important concept is the rate–distortion function. In a classical communication theoretic setup, this is defined as the infimum of rates R such that (R, D) is achievable. A similar definition of rate–distortion function in a game theoretic setup yields

$$R(D_e, D_d) \triangleq \inf\{R \mid (R, D_e, D_d) \text{ is achievable}\}. \tag{17}$$

By using the result of Theorem 21, we can upper bound the rate–distortion function for the Gaussian case. Towards that goal, we use the following result for the team theoretic setup where the distortion D is identical at the encoder and decoder as there is no bias at the encoder.

Lemma 23 (*Cover & Thomas, 2006, Theorem 10.3.2*). *Suppose that $b = 0$ in Problem 20. Consider i.i.d. Gaussian sources with a variance of σ^2. The rate distortion function for such a setup is given by $R(D) = \frac{1}{2} \log_2 \frac{\sigma^2}{D}$ if $0 \leq D \leq \sigma^2$ and $R(D) = 0$ if $D > \sigma^2$.*

Next, the following theorem, whose proof appears in Appendix K, presents our result for the game theoretic setup with a biased encoder.

Theorem 24. *Consider the multi-dimensional cheap talk problem with i.i.d. Gaussian sources where the bias term b is the same at each dimension. The rate–distortion function for such a setup satisfies $R(D_e, D_d) \leq \frac{1}{2} \log_2 \frac{\sigma^2}{\min\{D_d, D_e - b^2\}}$ if $0 \leq \min\{D_d, D_e - b^2\} \leq \sigma^2$, and $R(D_e, D_d) = 0$ if $\min\{D_d, D_e - b^2\} > \sigma^2$.*

Remark 25. A related result can be found in Kazıklı et al. (2022, Theorem 8). It is shown that having more bins in the quantized encoding policy leads to reduced distortion values if the scalar source has a log-concave distribution. In other words, a large rate leads to smaller expected costs for both players under a log-concave source assumption, which holds for the Gaussian case.

Remark 26. For multi-dimensional i.i.d. Gaussian sources, there exists a Nash equilibrium where the encoder reveals $(n-1)$ dimensions and applies a signaling game policy for the remaining dimension X_n with an arbitrary number of bins. Since the

Gaussian distribution is log-concave, from Kazıklı et al. (2022, Theorem 8), it follows that the expected costs of both players reduce when the number of bins for the quantization policy applied to X_n is increased. In addition, it is also possible to have a Nash equilibrium with infinitely many bins applied to X_n due to Kazıklı et al. (2022, Theorem 13). Hence, a Nash equilibrium where $X_1, \ldots, X_{n-1}$ are revealed and a quantization policy with infinitely many bins applied to X_n corresponds to a payoff dominant Nash equilibrium (Harsanyi & Selten, 1988).

5. Conclusion

We have analyzed a quadratic multi-dimensional cheap talk problem. First, we have derived the necessary general conditions for a Nash equilibrium considering any joint source distribution. In particular, we have shown that decoder actions at a Nash equilibrium need to satisfy a geometric condition that essentially prevents any two decoder actions from being arbitrarily close to each other depending on their difference as vectors and the bias vector. Then, we have investigated continuum of decoder actions considering two-dimensional sources and provided a condition that a continuum of decoder actions must satisfy in any Nash equilibrium. Then, we have derived necessary and sufficient conditions under which a linear Nash equilibrium exists considering i.i.d. sources. These conditions require a Gaussian or a symmetric source density. Moreover, we have formulated a rate–distortion theoretic problem for the cheap talk setup and have solved the Gaussian case.

Acknowledgments

Appendix A. Proof of Lemma 7

(i) Let there be two bins, $\mathcal{B}^\alpha$ and $\mathcal{B}^\beta$. Denote their centroids by $\boldsymbol{u}^\alpha = \mathbb{E}[\boldsymbol{M}|\boldsymbol{M} \in \mathcal{B}^\alpha]$ and $\boldsymbol{u}^\beta = \mathbb{E}[\boldsymbol{M}|\boldsymbol{M} \in \mathcal{B}^\beta]$. The encoder is indifferent between the decoder actions $\boldsymbol{u}^\alpha$ and $\boldsymbol{u}^\beta$ for source observation values $\bar{\boldsymbol{m}}$ which satisfy the following:

$$c^e(\bar{\boldsymbol{m}}, \boldsymbol{u}^\alpha) = c^e(\bar{\boldsymbol{m}}, \boldsymbol{u}^\beta)$$
$$\Leftrightarrow \|\bar{\boldsymbol{m}} - \boldsymbol{u}^\alpha - \boldsymbol{b}\|^2 = \|\bar{\boldsymbol{m}} - \boldsymbol{u}^\beta - \boldsymbol{b}\|^2$$
$$\Leftrightarrow (2\bar{\boldsymbol{m}} - (\boldsymbol{u}^\beta + \boldsymbol{u}^\alpha + 2\boldsymbol{b}))^T(\boldsymbol{u}^\beta - \boldsymbol{u}^\alpha) = 0. \quad (18)$$

In other words, if an observation satisfies (18), the encoder's costs are the same under the decoder actions $\boldsymbol{u}^\alpha$ and $\boldsymbol{u}^\beta$. From (18), it is seen that these $\bar{\boldsymbol{m}}$ values define a hyperplane orthogonal to $(\boldsymbol{u}^\beta - \boldsymbol{u}^\alpha)$. Given any source observation $\boldsymbol{m} = \bar{\boldsymbol{m}} + \Delta(\boldsymbol{u}^\beta - \boldsymbol{u}^\alpha)$ with $\Delta > 0$ where $\bar{\boldsymbol{m}}$ satisfies (18), the encoder prefers the decoder action $\boldsymbol{u}^\beta$ over the decoder action $\boldsymbol{u}^\alpha$ since the following holds:

$$c^e(\bar{\boldsymbol{m}} + \Delta(\boldsymbol{u}^\beta - \boldsymbol{u}^\alpha), \boldsymbol{u}^\beta) - c^e(\bar{\boldsymbol{m}} + \Delta(\boldsymbol{u}^\beta - \boldsymbol{u}^\alpha), \boldsymbol{u}^\alpha)$$
$$= -2\Delta\|\boldsymbol{u}^\beta - \boldsymbol{u}^\alpha\|^2 < 0. \quad (19)$$

This implies that $\mathcal{B}^\alpha$ and $\mathcal{H}_1$ are disjoint sets where

$$\mathcal{H}_1 \triangleq \{\boldsymbol{m} \mid \boldsymbol{m} = \bar{\boldsymbol{m}} + \Delta(\boldsymbol{u}^\beta - \boldsymbol{u}^\alpha)$$
$$\text{where } \bar{\boldsymbol{m}} \text{ satisfies (18) and } \Delta > 0\}. \quad (20)$$

Similarly, given any source observation $\boldsymbol{m} = \bar{\boldsymbol{m}} + \Delta(\boldsymbol{u}^\beta - \boldsymbol{u}^\alpha)$ with $\Delta < 0$, the encoder prefers the decoder action $\boldsymbol{u}^\alpha$ over the decoder action $\boldsymbol{u}^\beta$. It follows that $\mathcal{B}^\beta$ and $\mathcal{H}_2$ are disjoint sets where

$$\mathcal{H}_2 \triangleq \{\boldsymbol{m} \mid \boldsymbol{m} = \bar{\boldsymbol{m}} + \Delta(\boldsymbol{u}^\beta - \boldsymbol{u}^\alpha)$$
$$\text{where } \bar{\boldsymbol{m}} \text{ satisfies (18) and } \Delta < 0\}. \quad (21)$$

Furthermore, the plane specified by (18) intersects the affine set $\lambda\boldsymbol{u}^\beta + (1-\lambda)\boldsymbol{u}^\alpha$ with $\lambda \in \mathbb{R}$ at a single point due to the fact that $(\boldsymbol{u}^\beta - \boldsymbol{u}^\alpha)$ and the hyperplane specified by (18) are orthogonal. In order to find the value of λ that gives this intersection point, we solve (18) and $\bar{\boldsymbol{m}} = \bar{\lambda}\boldsymbol{u}^\beta + (1 - \bar{\lambda})\boldsymbol{u}^\alpha$ together and obtain an expression for $\bar{\lambda}$ in the following:

$$(2\bar{\boldsymbol{m}} - \boldsymbol{u}^\beta - \boldsymbol{u}^\alpha - 2\boldsymbol{b})^T(\boldsymbol{u}^\beta - \boldsymbol{u}^\alpha) = 0$$
$$\Leftrightarrow (2\bar{\lambda} - 1) = \frac{2(\boldsymbol{u}^\beta - \boldsymbol{u}^\alpha)^T\boldsymbol{b}}{\|\boldsymbol{u}^\beta - \boldsymbol{u}^\alpha\|^2}. \quad (22)$$

We know that the centroid conditions require $\boldsymbol{u}^\alpha = \mathbb{E}[\boldsymbol{M}|\boldsymbol{M} \in \mathcal{B}^\alpha]$ and $\boldsymbol{u}^\beta = \mathbb{E}[\boldsymbol{M}|\boldsymbol{M} \in \mathcal{B}^\beta]$. Since $\mathcal{B}^\alpha$ and $\mathcal{H}_1$ are disjoint sets, and $\mathcal{B}^\beta$ and $\mathcal{H}_2$ are disjoint sets, we need $0 \le \bar{\lambda} \le 1$, which is equivalent to $|2\bar{\lambda} - 1| \le 1$. By combining this inequality with (22), it follows that (4) holds.

(ii) As noted above, $\mathcal{B}^\alpha$ and $\mathcal{B}^\beta$ do not intersect with $\mathcal{H}_1$ and $\mathcal{H}_2$, respectively, where $\mathcal{H}_1$ and $\mathcal{H}_2$ are defined in (20) and (21). These regions lead to the decomposition specified by the hyperplane in (5), which is orthogonal to $(\boldsymbol{u}^\beta - \boldsymbol{u}^\alpha)$.

(iii) The decision regions for two decoder actions must be constructed by computing and intersecting half spaces. Since a half space is a convex set and intersection operation preserves convexity, the quantization bins must be convex (Boyd & Vandenberghe, 2004, p. 36).

Appendix B. Proof of Lemma 9

For the cost function of the encoder, we can write

$$c^e(\boldsymbol{m}, \boldsymbol{u}) = (\boldsymbol{m} - \boldsymbol{u} - \boldsymbol{b})^T(\boldsymbol{m} - \boldsymbol{u} - \boldsymbol{b})$$
$$= (\mathcal{T}^{-1}\mathcal{T}(\boldsymbol{m} - \boldsymbol{u} - \boldsymbol{b}))^T(\mathcal{T}^{-1}\mathcal{T}(\boldsymbol{m} - \boldsymbol{u} - \boldsymbol{b}))$$
$$= (\mathcal{T}(\boldsymbol{m} - \boldsymbol{u} - \boldsymbol{b}))^T(\mathcal{T}^{-1})^T(\mathcal{T}^{-1})(\mathcal{T}(\boldsymbol{m} - \boldsymbol{u} - \boldsymbol{b}))$$
$$= \tilde{b}^{-1}(\boldsymbol{x} - \boldsymbol{y} - [0, \tilde{b}]^T)^T(\boldsymbol{x} - \boldsymbol{y} - [0, \tilde{b}]^T)$$
$$= c_t^e(\boldsymbol{x}, \boldsymbol{y})\tilde{b}^{-1}, \quad (23)$$

where $\tilde{b}$ is specified in (13), and the fourth equation uses $(\mathcal{T}^{-1})^T(\mathcal{T}^{-1}) = \tilde{b}^{-1}I$ with I denoting identity matrix, $\mathcal{T}\boldsymbol{b} = [0, \tilde{b}]^T$, $\mathcal{T}\boldsymbol{m} = \boldsymbol{x}$ and $\mathcal{T}\boldsymbol{u} = \boldsymbol{y}$. In a similar manner, the cost function of the decoder can be expressed as

$$c^d(\boldsymbol{m}, \boldsymbol{u}) = (\boldsymbol{m} - \boldsymbol{u})^T(\boldsymbol{m} - \boldsymbol{u})$$
$$= (\mathcal{T}(\boldsymbol{m} - \boldsymbol{u}))^T(\mathcal{T}^{-1})^T(\mathcal{T}^{-1})(\mathcal{T}(\boldsymbol{m} - \boldsymbol{u}))$$
$$= \tilde{b}^{-1}(\boldsymbol{x} - \boldsymbol{y})^T(\boldsymbol{x} - \boldsymbol{y}) = c_t^d(\boldsymbol{x}, \boldsymbol{y})\tilde{b}^{-1}, \quad (24)$$

where the third equation uses $(\mathcal{T}^{-1})^T(\mathcal{T}^{-1}) = \tilde{b}^{-1}I$, $\mathcal{T}\boldsymbol{m} = \boldsymbol{x}$ and $\mathcal{T}\boldsymbol{u} = \boldsymbol{y}$. Note that the factor of $(1/\tilde{b})$ is canceled in the definitions of $c_t^e(\boldsymbol{x}, \boldsymbol{y})$ and $c_t^d(\boldsymbol{x}, \boldsymbol{y})$ for notational convenience. Since $\tilde{b}$ is a constant and $\tilde{b} > 0$, this cancellation does not change the goals of the players.

Note that a fixed and invertible transformation of observation and action variables independent of the policies $\tilde{\gamma}^e(\cdot)$ and $\tilde{\gamma}^d(\cdot)$ is considered. Moreover, since the transformation is invertible, there is no loss of information at the encoder and the decoder due to the transformation. If $\gamma^{*,e}(\cdot)$ and $\gamma^{*,d}(\cdot)$ satisfy (3), then $\tilde{\gamma}^{*,e}(\cdot) = \gamma^{*,e}(\mathcal{T}^{-1}(\cdot))$ and $\tilde{\gamma}^{*,d}(\cdot) = \mathcal{T}(\gamma^{*,d}(\cdot))$ satisfy $\tilde{J}^e(\tilde{\gamma}^{*,e}, \tilde{\gamma}^{*,d}) \le \tilde{J}^e(\tilde{\gamma}^e, \tilde{\gamma}^{*,d})$ for all $\tilde{\gamma}^e$ and $\tilde{J}^d(\tilde{\gamma}^{*,e}, \tilde{\gamma}^{*,d}) \le \tilde{J}^d(\tilde{\gamma}^{*,e}, \tilde{\gamma}^d)$ for all $\tilde{\gamma}^d$ since the cost functions in the transformed and original formulations are essentially the same as shown in (23) and (24). Similarly, if $\tilde{\gamma}^{*,e}(\cdot)$ and $\tilde{\gamma}^{*,d}(\cdot)$ satisfy $\tilde{J}^e(\tilde{\gamma}^{*,e}, \tilde{\gamma}^{*,d}) \le \tilde{J}^e(\tilde{\gamma}^e, \tilde{\gamma}^{*,d})$ for all $\tilde{\gamma}^e$ and $\tilde{J}^d(\tilde{\gamma}^{*,e}, \tilde{\gamma}^{*,d}) \le \tilde{J}^d(\tilde{\gamma}^{*,e}, \tilde{\gamma}^d)$ for all $\tilde{\gamma}^d$,

then $\gamma^{*,e}(\cdot) = \tilde{\gamma}^{*,e}(\mathcal{T}(\cdot))$ and $\gamma^{*,d}(\cdot) = \mathcal{T}^{-1}(\tilde{\gamma}^{*,d}(\cdot))$ satisfy (3). These reveal that an equilibrium under the proposed (original) formulation is also an equilibrium under the original (proposed) formulation. This equivalence can be viewed as a special case of the result in Sanjari, Başar, and Yüksel (2021, Theorem 3.1) where it is shown that for a dynamic stochastic game setup one can equivalently consider its policy-independent static reduction under the Nash equilibrium concept.

Appendix C. Proof of Lemma 10

For a given encoding policy, the aim of the decoder is to minimize $\tilde{J}^d(\tilde{\gamma}^e, \tilde{\gamma}^d) = \mathbb{E}[c_t^d(\mathbf{X}, \mathbf{Y})]$ where $c_t^d(\cdot, \cdot)$ is as in (24). Since the expression in (24) involves a sum of squared error terms, the result immediately follows.

Appendix D. Proof of Lemma 11

This result is a consequence of Lemma 7. By using the cost function of the encoder and the decoder considering the equivalent formulation, it can be shown that the condition in (4) translates to the condition that for any decoder actions $\mathbf{y}^\alpha = [y_1^\alpha, y_2^\alpha]^T$ and $\mathbf{y}^\beta = [y_1^\beta, y_2^\beta]^T$, we have that

$$0 \leq (y_1^\alpha - y_1^\beta)^2 + (y_2^\alpha - y_2^\beta)^2 - 2\tilde{b}|y_2^\alpha - y_2^\beta| \triangleq g(\mathbf{y}^\alpha, \mathbf{y}^\beta). \quad (25)$$

Let there be a continuum of decoder actions that does not have a constant y_2 coordinate. Since a continuum of decoder actions with a constant y_1 coordinate is not allowed due to (25), it is possible to partition a continuum of actions so that it consists of representations of the form $\mathbf{y} = [y_1, h_i(y_1)]^T$ for some continuous functions $h_i(\cdot)$ with $i \in \{1, \dots, k\}$. Here, we represent the second coordinate of the continuum as a function of the first coordinate. In the following, we take such a continuum denoted by $\mathbf{y} = [y_1, h(y_1)]^T$ for some continuous function $h(\cdot)$ and prove that it leads to a contradiction when $h(\cdot)$ is not a constant function.

Suppose, by contradiction, that $h(\cdot)$ is not a constant function. It follows that we can find two decoder actions $\mathbf{y}^\alpha = [y_1^\alpha, y_2^\alpha]^T$ and $\mathbf{y}^\beta = [y_1^\beta, y_2^\beta]^T$ on the continuum with the property that $y_2^\alpha \neq y_2^\beta$ and that $y_1^\alpha < y_1^\beta$. Without loss of generality, take $y_2^\alpha < y_2^\beta$. In the following, we first prove the result by making the additional assumption that $h(\cdot)$ is differentiable. By this assumption, we are able to invoke the standard mean value theorem of calculus to conclude the result. However, the result holds also for a non-differentiable $h(\cdot)$. We first prove the result for a differentiable $h(\cdot)$ since this proof is more intuitive. Then, we prove the result in the general case when $h(\cdot)$ is non-differentiable.

Let $h(\cdot)$ be differentiable. By the mean value theorem of calculus, there exists $y_1^\gamma \in [y_1^\alpha, y_1^\beta]$ such that $h'(y_1^\gamma) = (y_2^\beta - y_2^\alpha)/(y_1^\beta - y_1^\alpha) > 0$. In particular, we can find a decoder action $\mathbf{y}^\gamma$ on the continuum with $y_1^\gamma \in [y_1^\alpha, y_1^\beta]$ and $y_2^\gamma = h(y_1^\gamma)$ such that the derivative of $h(\cdot)$ at y_1^γ is strictly positive. Next, we take another decoder action $\mathbf{y}^\eta = [y_1^\eta, h(y_1^\eta)]^T$ on the continuum and vary its first coordinate to reach a contradiction to (25). In particular, if we express the condition imposed by (25) for the decoder action $\mathbf{y}^\gamma$ and a decoder action on the continuum denoted by $\mathbf{y}^\eta$, we get

$$g(\mathbf{y}^\eta, \mathbf{y}^\gamma) = (y_1^\eta - y_1^\gamma)^2$$
$$+ (h(y_1^\eta) - h(y_1^\gamma))^2 - 2\tilde{b}|h(y_1^\eta) - h(y_1^\gamma)| \geq 0.$$

When $y_1^\eta > y_1^\gamma$, and $(y_1^\eta - y_1^\gamma)$ is sufficiently small, we have $h(y_1^\eta) > h(y_1^\gamma)$ due to a positive derivative at y_1^γ. For fixed y_1^γ, if we take the derivative of $g(\mathbf{y}^\eta, \mathbf{y}^\gamma)$ with respect to y_1^η, we get

$$\frac{dg(\mathbf{y}^\eta, \mathbf{y}^\gamma)}{dy_1^\eta} = 2(y_1^\eta - y_1^\gamma)$$
$$+ 2(h(y_1^\eta) - h(y_1^\gamma))h'(y_1^\eta) - 2\tilde{b}h'(y_1^\eta). \quad (26)$$

If we take $y_1^\eta = y_1^\gamma$ in (26), then the first two terms are zero while the third term is negative since $h'(y_1^\gamma) > 0$ and $\tilde{b} > 0$. Therefore, we have that $\frac{dg(\mathbf{y}^\eta, \mathbf{y}^\gamma)}{dy_1^\eta}\Big|_{y_1^\eta = y_1^\gamma} < 0$. This is a contradiction to $g(\mathbf{y}^\eta, \mathbf{y}^\gamma) \geq 0$ since $g(\mathbf{y}^\gamma, \mathbf{y}^\gamma) = 0$ and $\frac{dg(\mathbf{y}^\eta, \mathbf{y}^\gamma)}{dy_1^\eta}\Big|_{y_1^\eta = y_1^\gamma} < 0$.

Therefore, it must be that $h(\cdot)$ is a constant function, which means that a continuum of actions must have a constant y_2 coordinate.

Now, consider the general case when $h(\cdot)$ is not differentiable. By the mean value theorem in Hiriart-Urruty (1980, Corollary 1), there exists $y_1^* \in [y_1^\alpha, y_1^\beta]$ such that either $s^* \in \partial h(y_1^*)$ or $s^* \in -\partial(-h(y_1^*))$ holds where $s^* \triangleq (y_2^\beta - y_2^\alpha)/(y_1^\beta - y_1^\alpha) > 0$, and $\partial h(y_1^*)$ denotes the set of subgradients of $h(\cdot)$ at y_1^*. We take two decoder actions and employ the condition in (25) to reach a contradiction. Let $\mathbf{y}^\gamma$ and $\mathbf{y}^\eta$ denote these two decoder actions, which are expressed as $y_1^\gamma = y_1$, $y_2^\gamma = h(y_1^\gamma)$, $y_1^\eta = y_1 + td'$ and $y_2^\eta = h(y_1^\eta)$. If we express the condition imposed by (25) for the decoder actions $\mathbf{y}^\gamma$ and $\mathbf{y}^\eta$, we get

$$g(\mathbf{y}^\eta, \mathbf{y}^\gamma)$$
$$= (y_1^\eta - y_1^\gamma)^2 + (h(y_1^\eta) - h(y_1^\gamma))^2 - 2\tilde{b}|h(y_1^\eta) - h(y_1^\gamma)|$$
$$= (td')^2 + (h(y_1 + td') - h(y_1))^2$$
$$- 2\tilde{b}|h(y_1 + td') - h(y_1)| \geq 0. \quad (27)$$

If t and d' are positive and sufficiently small, the inequality in (27) can be expressed as

$$|h(y_1 + td') - h(y_1)| + \sqrt{\tilde{b}^2 - (td')^2} - \tilde{b} \leq 0. \quad (28)$$

Since t is positive, it follows from (28) that

$$\frac{|h(y_1 + td') - h(y_1)|}{t} + \frac{\sqrt{\tilde{b}^2 - (td')^2} - \sqrt{\tilde{b}^2}}{t} \leq 0. \quad (29)$$

Thus, (29) implies that

$$\limsup_{\substack{y_1 \to y_1^* \\ t \downarrow 0}} \inf_{d' \to d} \frac{|h(y_1 + td') - h(y_1)|}{t}$$
$$+ \frac{\sqrt{\tilde{b}^2 - (td')^2} - \sqrt{\tilde{b}^2}}{t} \leq 0, \quad (30)$$

where we take $d > 0$ small so that the assumption of having a small d' holds. Since the limit of the second term in (30) is zero, we get

$$\limsup_{\substack{y_1 \to y_1^* \\ t \downarrow 0}} \inf_{d' \to d} \frac{|h(y_1 + td') - h(y_1)|}{t} = 0. \quad (31)$$

However, this is a contradiction to the fact that either $s^* \in \partial h(y_1^*)$ or $s^* \in -\partial(-h(y_1^*))$ holds.

Appendix E. Proof of Lemma 12

In order to prove this result, we assume that there exists a continuum of decoder actions whose support does not extend to the boundaries of the support and then reach a contradiction. Let $\mathbf{y}^\alpha = [y_1^\alpha, y_2^\alpha]^T$ and $\mathbf{y}^\beta = [y_1^\beta, y_2^\beta]^T$ be decoder actions on the continuum with $y_1^\alpha < y_1^\beta$ and $y_2^\alpha = y_2^\beta$ such that there exist decoder actions for all values of y_1 satisfying $y_1^\alpha \leq y_1 \leq y_1^\beta$ and $y_2 = y_2^\alpha$ where $x_1^L(y_2^\alpha) < y_1^\alpha$ and $y_1^\beta < x_1^U(y_2^\beta)$. As mentioned earlier, $x_1^L(x_2)$ and $x_1^U(x_2)$ respectively denote the lower and upper boundaries of the support for X_1 given that $X_2 = x_2$. In addition, suppose that there exists $\delta > 0$ such that there is no decoder action with $y_1^\beta < y_1 < y_1^\beta + \delta$ and $y_2 = y_2^\beta$. Similarly, suppose

that there exists $\tilde{\delta} > 0$ such that there is no decoder action with $y_1^\alpha - \tilde{\delta} < y_1 < y_1^\alpha$ and $y_2 = y_2^\alpha$. In the following, we focus on the decoder action $\boldsymbol{y}^\beta$ and show that the assumption of $y_1^\beta < x_1^U(y_2^\beta)$ leads to a contradiction. A similar approach can be taken for the decoder action $\boldsymbol{y}^\alpha$ to prove that the assumption of $x_1^L(y_2^\alpha) < y_1^\alpha$ leads to a contradiction.

Let $\mathcal{B}^\beta$ denote the bin corresponding to the decoder action $\boldsymbol{y}^\beta$, i.e., $\boldsymbol{y}^\beta = \mathbb{E}[\boldsymbol{X}|\boldsymbol{X} \in \mathcal{B}^\beta]$. We will obtain a contradiction that $\mathcal{B}^\beta$ contains observations with $x_1 > y_1^\beta$, whereas it does not contain any observation with $x_1 < y_1^\beta$. This is a contradiction to $\boldsymbol{y}^\beta = \mathbb{E}[\boldsymbol{X}|\boldsymbol{X} \in \mathcal{B}^\beta]$ since we assume that every non-empty open set has a positive probability measure in Assumption 2.

We first show that $\mathcal{B}^\beta$ does not contain any observations with $x_1 < y_1^\beta$. Take an observation $\boldsymbol{x}$ with $x_1 < y_1^\beta$. Since the continuum of decoder actions is supported on $y_1 \in [y_1^\alpha, y_1^\beta]$ and $y_2 = y_2^\beta$, we can find a decoder action $\boldsymbol{y}^\nu$ on the continuum with coordinates $y_1^\nu = \max\{x_1, y_1^\alpha\}$ and $y_2^\nu = y_2^\beta$. For this decoder action, when $x_1 < y_1^\beta$, we have that $c_t^e(\boldsymbol{x}, \boldsymbol{y}^\nu) < c_t^e(\boldsymbol{x}, \boldsymbol{y}^\beta)$. This implies that any observation with $x_1 < y_1^\beta$ cannot be an element of $\mathcal{B}^\beta$.

Next, we prove that $\mathcal{B}^\beta$ must contain observations with $x_1 > y_1^\beta$ under the assumed configuration, which, however, contradicts $\boldsymbol{y}^\beta = \mathbb{E}[\boldsymbol{X}|\boldsymbol{X} \in \mathcal{B}^\beta]$. While proving that $\mathcal{B}^\beta$ must contain observations with $x_1 > y_1^\beta$, we use the result from Lemma 7 which imposes the condition in (25) in the transformed coordinate system considering any two decoder actions at a Nash equilibrium. In particular, for any decoder action $\boldsymbol{y}^\eta$, we need $g(\boldsymbol{y}^\beta, \boldsymbol{y}^\eta) \geq 0$ where $g(\cdot, \cdot)$ is defined in (25). Note that other decoder actions impose additional constraints on the region where decoder actions can exist at a Nash equilibrium. Nonetheless, for our purpose, it suffices to use the condition $g(\boldsymbol{y}^\beta, \boldsymbol{y}^\eta) \geq 0$. Our aim is to show that for all $\boldsymbol{y}^\eta$ satisfying $g(\boldsymbol{y}^\beta, \boldsymbol{y}^\eta) \geq 0$, it is possible to find a region of observations with $x_1 > y_1^\beta$ where $c_t^e(\boldsymbol{x}, \boldsymbol{y}^\beta) < c_t^e(\boldsymbol{x}, \boldsymbol{y}^\eta)$. This implies that the intersection of these observation regions must be the bin for $\boldsymbol{y}^\beta$, i.e., $\mathcal{B}^\beta$. To conclude the result, we will show that this intersection is not empty. Note that $c_t^e(\boldsymbol{x}, \boldsymbol{y}^\beta) < c_t^e(\boldsymbol{x}, \boldsymbol{y}^\eta)$ is equivalent to

$$(y_1^\eta - y_1^\beta)(y_1^\eta + y_1^\beta - 2x_1)$$
$$+ (y_2^\eta - y_2^\beta)(y_2^\eta + y_2^\beta + 2\tilde{b} - 2x_2) > 0. \tag{32}$$

We will use this equivalent expression in the remainder of the proof.

When $y_2^\eta = y_2^\beta$, due to the discontinuity assumption at $\boldsymbol{y}^\beta$, for any observation with $y_1^\beta < x_1 < y_1^\beta + \delta/2$, the inequality $c_t^e(\boldsymbol{x}, \boldsymbol{y}^\beta) < c_t^e(\boldsymbol{x}, \boldsymbol{y}^\eta)$ holds. Therefore, it is sufficient to consider decoder actions with $y_2^\eta \neq y_2^\beta$. We can express a decoder action $\boldsymbol{y}^\eta$ satisfying $y_2^\eta \neq y_2^\beta$ and $g(\boldsymbol{y}^\beta, \boldsymbol{y}^\eta) \geq 0$ as $\boldsymbol{y}^\eta = \boldsymbol{y}^\beta + \lambda(\boldsymbol{y}^\gamma - \boldsymbol{y}^\beta)$ for some $\lambda \geq 1$ where $\boldsymbol{y}^\gamma$ satisfies $g(\boldsymbol{y}^\beta, \boldsymbol{y}^\gamma) = 0$. In particular, we have

$$(y_1^\eta - y_1^\beta)^2 + (y_2^\eta - y_2^\beta)^2 - 2\tilde{b}|y_2^\eta - y_2^\beta|$$
$$= \lambda^2(y_1^\gamma - y_1^\beta)^2 + \lambda^2(y_2^\gamma - y_2^\beta)^2 - 2|\lambda| \, |y_2^\gamma - y_2^\beta|\tilde{b}$$
$$\overset{(a)}{=} \lambda^2(2\tilde{b}|y_2^\gamma - y_2^\beta|) - 2|\lambda| \, |y_2^\gamma - y_2^\beta|\tilde{b}$$
$$= (2\tilde{b}|y_2^\gamma - y_2^\beta|)(\lambda^2 - |\lambda|)$$

where (a) follows from $g(\boldsymbol{y}^\beta, \boldsymbol{y}^\gamma) = 0$, and the final expression implies that we can write $\boldsymbol{y}^\eta = \boldsymbol{y}^\beta + \lambda(\boldsymbol{y}^\gamma - \boldsymbol{y}^\beta)$ with $\lambda \geq 1$ so that $g(\boldsymbol{y}^\beta, \boldsymbol{y}^\eta) \geq 0$ holds. It is seen that when $\lambda > 1$, for a given observation $\boldsymbol{x}$ satisfying $c_t^e(\boldsymbol{x}, \boldsymbol{y}^\beta) < c_t^e(\boldsymbol{x}, \boldsymbol{y}^\gamma)$, the following holds:

$$(y_1^\eta - y_1^\gamma)(y_1^\eta + y_1^\gamma - 2x_1)$$
$$+ (y_2^\eta - y_2^\gamma)(y_2^\eta + y_2^\gamma + 2\tilde{b} - 2x_2)$$
$$= (\lambda - 1)(y_1^\gamma - y_1^\beta)((\lambda + 1)y_1^\gamma + (1 - \lambda)y_1^\beta - 2x_1)$$
$$+ (\lambda - 1)(y_2^\gamma - y_2^\beta)((\lambda + 1)y_2^\gamma + (1 - \lambda)y_2^\beta + 2\tilde{b} - 2x_2)$$

$$> (\lambda - 1)(y_1^\gamma - y_1^\beta)((\lambda + 1)y_1^\gamma + (1 - \lambda)y_1^\beta - y_1^\gamma - y_1^\beta)$$
$$+ (\lambda - 1)(y_2^\gamma - y_2^\beta)((\lambda + 1)y_2^\gamma + (1 - \lambda)y_2^\beta - y_2^\gamma - y_2^\beta)$$
$$= \lambda(\lambda - 1)((y_1^\gamma - y_1^\beta)^2 + (y_2^\gamma - y_2^\beta)^2) > 0,$$

where the first inequality is due to $c_t^e(\boldsymbol{x}, \boldsymbol{y}^\beta) < c_t^e(\boldsymbol{x}, \boldsymbol{y}^\gamma)$ and $\lambda > 1$, and the last inequality follows from $\boldsymbol{y}^\beta \neq \boldsymbol{y}^\gamma$ and $\lambda > 1$. This shows that $c_t^e(\boldsymbol{x}, \boldsymbol{y}^\beta) < c_t^e(\boldsymbol{x}, \boldsymbol{y}^\gamma)$ implies $c_t^e(\boldsymbol{x}, \boldsymbol{y}^\gamma) < c_t^e(\boldsymbol{x}, \boldsymbol{y}^\eta)$. Hence, it is sufficient to consider decoder actions $\boldsymbol{y}^\gamma$ with $g(\boldsymbol{y}^\beta, \boldsymbol{y}^\gamma) = 0$.

Next, we show that for any decoder action $\boldsymbol{y}^\gamma$ satisfying $g(\boldsymbol{y}^\beta, \boldsymbol{y}^\gamma) = 0$, there exists a nonempty region of observations with $c_t^e(\boldsymbol{x}, \boldsymbol{y}^\beta) < c_t^e(\boldsymbol{x}, \boldsymbol{y}^\gamma)$ and $x_1 > y_1^\beta$. Towards that goal, we consider different cases and treat each of these cases separately. In the following, we take $y_2^\beta < x_2 < y_2^\beta + 2\tilde{b}$ and specify a nontrivial interval for x_1 so that the resulting observation satisfies the desired property.

(i) Let $\boldsymbol{y}^\gamma$ satisfy $y_2^\beta < y_2^\gamma < y_2^\beta + 2\tilde{b}$ and $y_1^\gamma = y_1^\beta + ((y_2^\gamma - y_2^\beta)(y_2^\gamma - y_2^\beta + 2\tilde{b}))^{1/2}$. Let

$$y_1^\beta < x_1 < y_1^\beta + \frac{(y_2^\gamma - y_2^\beta)(y_2^\beta - x_2 + 2\tilde{b})}{(y_1^\gamma - y_1^\beta)}. \tag{33}$$

After some manipulations, it can be shown that (32) and equivalently $c_t^e(\boldsymbol{x}, \boldsymbol{y}^\beta) < c_t^e(\boldsymbol{x}, \boldsymbol{y}^\gamma)$ hold for any observation that satisfies (33) and $y_2^\beta < x_2 < y_2^\beta + 2\tilde{b}$. Notice also that if (33) holds for some y_2^γ, then it automatically holds for any $\tilde{y}_2^\gamma$ satisfying $y_2^\gamma < \tilde{y}_2^\gamma < y_2^\beta + 2\tilde{b}$. Therefore, multiple decoder actions satisfy the assumptions of this case, it is sufficient to consider the one with the minimum y_2 coordinate.

(ii) Let $\boldsymbol{y}^\gamma$ be such that $y_2^\beta < y_2^\gamma \leq y_2^\beta + 2\tilde{b}$ and $y_1^\gamma = y_1^\beta - ((y_2^\gamma - y_2^\beta)(y_2^\gamma - y_2^\beta + 2\tilde{b}))^{1/2}$ hold. In this case, one can show that for any observation with $y_1^\beta < x_1$ and $y_2^\beta < x_2 < y_2^\beta + 2\tilde{b}$, the inequality $c_t^e(\boldsymbol{x}, \boldsymbol{y}^\beta) < c_t^e(\boldsymbol{x}, \boldsymbol{y}^\gamma)$ holds.

(iii) Let $\boldsymbol{y}^\gamma$ satisfy $y_2^\beta - 2\tilde{b} < y_2^\gamma < y_2^\beta$ and $y_1^\gamma = y_1^\beta + ((y_2^\gamma - y_2^\beta)(y_2^\gamma - y_2^\beta - 2\tilde{b}))^{1/2}$. Let

$$y_1^\beta < x_1 < y_1^\beta + \frac{(y_2^\beta - y_2^\gamma)(x_2 - y_2^\beta)}{(y_1^\gamma - y_1^\beta)}. \tag{34}$$

After some manipulations, it can be shown that (32) and equivalently $c_t^e(\boldsymbol{x}, \boldsymbol{y}^\beta) < c_t^e(\boldsymbol{x}, \boldsymbol{y}^\gamma)$ hold for any observation that satisfies (34) and $y_2^\beta < x_2 < y_2^\beta + 2\tilde{b}$. In addition, notice that if (34) holds for some y_2^γ, then it automatically holds for any $\tilde{y}_2^\gamma$ satisfying $y_2^\gamma - 2\tilde{b} < \tilde{y}_2^\gamma < y_2^\gamma$. Therefore, multiple decoder actions satisfy the assumptions of this case, it is sufficient to consider the one with the maximum y_2 coordinate.

(iv) Let $\boldsymbol{y}^\gamma$ be such that $y_2^\beta - 2\tilde{b} \leq y_2^\gamma < y_2^\beta$ and $y_1^\gamma = y_1^\beta - ((y_2^\gamma - y_2^\beta)(y_2^\gamma - y_2^\beta - 2\tilde{b}))^{1/2}$. In this case, one can show that for any observation with $y_1^\beta < x_1$ and $y_2^\beta < x_2 < y_2^\beta + 2\tilde{b}$, the inequality $c_t^e(\boldsymbol{x}, \boldsymbol{y}^\beta) < c_t^e(\boldsymbol{x}, \boldsymbol{y}^\gamma)$ holds.

As a result, for any decoder action $\boldsymbol{y}^\eta$ satisfying $g(\boldsymbol{y}^\beta, \boldsymbol{y}^\eta) \geq 0$ and for a given x_2 with $y_2^\beta < x_2 < y_2^\beta + 2\tilde{b}$, a nontrivial interval exists for x_1 with $x_1 > y_1^\beta$ such that $c_t^e(\boldsymbol{x}, \boldsymbol{y}^\beta) < c_t^e(\boldsymbol{x}, \boldsymbol{y}^\eta)$ holds. This implies that there exists a nonempty region of observations with $x_1 > y_1^\beta$ that must belong to $\mathcal{B}^\beta$. Fig. 6 illustrates this region of observations for an example scenario. Note that as long as $y_1^\beta < x_1^U(y_2^\beta)$, a nonempty subset of this observation region is in the support of the joint distribution. As a result, we obtain a contradiction to $\boldsymbol{y}^\beta = \mathbb{E}[\boldsymbol{X}|\boldsymbol{X} \in \mathcal{B}^\beta]$ since $\mathcal{B}^\beta$ does not contain any observation with $x_1 < y_1^\beta$, and $\mathcal{B}^\beta$ contains a nonempty region of observations with $x_1 > y_1^\beta$.
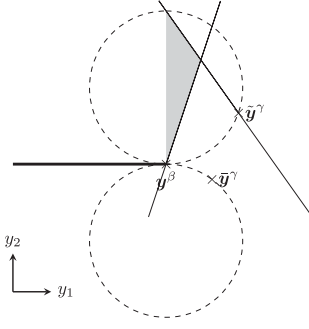
Fig. 6. Illustration of the proof technique employed in Lemma 12. Here, the horizontal solid line represents a continuum of decoder actions, and the decoder action on this continuum with the largest y_1 coordinate is denoted by $\boldsymbol{y}^\beta$. Inside the dashed circles, there cannot be a decoder action due to the condition imposed by Lemma 7 that for any $\boldsymbol{y}$, we have $g(\boldsymbol{y}^\beta, \boldsymbol{y}) \geq 0$ where $g(\cdot, \cdot)$ is defined in (25). If we place decoder actions $\bar{\boldsymbol{y}}^\gamma$ and $\tilde{\boldsymbol{y}}^\gamma$ on the dashed circles, then the shaded area must be the bin for $\boldsymbol{y}^\beta$, i.e., $\mathcal{B}^\beta$. However, the centroid of this shaded area cannot be $\boldsymbol{y}^\beta$, which is a contradiction.

Appendix F. Proof of Theorem 13

Suppose without loss of generality that $\mathbb{E}[M_1] = 0$. In the proof, we consider the equivalent formulation in Lemma 9.

(i) Since M_1 and M_2 are independent random variables, the problem decouples into two one-dimensional cheap talk problems where one of them involves an encoder with a zero bias. Hence, revealing the source corresponding to a zero bias leads to a linear Nash equilibrium.

(ii) If the source distribution is Gaussian, X_1 and X_2 are independent random variables. Thus, we obtain decoupled one-dimensional cheap talk problems where one of them involves an encoder with a zero bias. Then, revealing the random variable corresponding to a zero bias (i.e., X_1) yields an informative Nash equilibrium where the encoder is linear.

Now, suppose that the source distribution is not Gaussian. In this case, we show that there does not exist a Nash equilibrium with an encoding policy $z = \gamma^e(\boldsymbol{m}) = \alpha_1 m_2 - \alpha_2 m_1$ for any scalars α_1 and α_2. From Lemmas 11 and 12, we know that a continuum of actions must have a constant y_2 coordinate, say κ, and must be supported for all values of y_1 in the support of X_1 given that $X_2 = \kappa$. This implies that a necessary condition for a Nash equilibrium with an encoding policy $z = \gamma^e(\boldsymbol{m}) = \alpha_1 m_2 - \alpha_2 m_1$ is given by $\mathbb{E}[X_2|\bar{X} = \bar{x}] = \mathbb{E}[b_1 M_1 + b_2 M_2|\alpha_1 M_2 - \alpha_2 M_1 = \bar{x}] = 0$ for all $\bar{x}$ where $\bar{X} \triangleq \alpha_1 M_2 - \alpha_2 M_1$. If $\alpha_1 = 0$, $\alpha_2 = 0$, or $\frac{\alpha_1}{\alpha_2} \neq \frac{b_1}{b_2}$, one can decompose X_2 to show that the condition of $\mathbb{E}[X_2|\bar{X} = \bar{x}] = 0$ for all $\bar{x}$ is always violated. It follows that there does not exist a Nash equilibrium with a linear encoding policy $z = \gamma^e(\boldsymbol{m}) = \alpha_1 m_2 - \alpha_2 m_1$ when $\alpha_1 = 0$, $\alpha_2 = 0$, or $\frac{\alpha_1}{\alpha_2} \neq \frac{b_1}{b_2}$. It remains to investigate conditions under which an encoding policy $z = \gamma^e(\boldsymbol{m}) = b_1 m_2 - b_2 m_1$ leads to a Nash equilibrium.

If $b_1 \neq 0$, $b_2 \neq 0$ and $|b_1| \neq |b_2|$, the condition of $\mathbb{E}[b_1 M_1 + b_2 M_2|b_1 M_2 - b_2 M_1 = x_1] = \mathbb{E}[X_2|X_1 = x_1] = 0$ for all x_1 requires that the source distribution is Gaussian (Kagan et al., 1973, Theorem 5.3.1). Hence, if the encoder reveals $X_1 = b_1 M_2 - b_2 M_1$ completely without giving additional information, we obtain a single continuum that contains decoder actions with different second coordinates. Since this contradicts with Lemma 11, there cannot be a Nash equilibrium where the encoder conveys $X_1 = b_1 M_2 - b_2 M_1$ only. In addition, having more than one continuum

of decoder actions, each with a constant second coordinate, implies that $\mathbb{E}[X_2|X_1 = x_1] = 0$ for all x_1. Hence, it is not possible to have a Nash equilibrium with more than one continuum of decoder actions.

(iii) In the case of $b_1 = b_2$, the condition of $\mathbb{E}[X_2|X_1 = x_1] = 0$ for all x_1 requires that the source distribution is symmetric (almost everywhere) (Kagan et al., 1973, Theorem 5.3.1). It follows that if the source distribution is not symmetric, there does not exist an informative linear Nash equilibrium. Now, suppose that the source distribution is symmetric. Let the encoding policy be given by $z = \tilde{\gamma}^e(\boldsymbol{x}) = x_1$. In other words, the encoder reveals X_1 completely without giving any additional information. Then, the best response of the decoder yields a single continuum of decoder actions. In particular, there only exist decoder actions for all values of $y_1 \in [x_1^L(0), x_1^U(0)]$ and $y_2 = 0$ where $x_1^L(0)$ and $x_1^U(0)$ respectively denote lower and upper boundaries of the support for X_1 given that $X_2 = 0$. Now, we suppose that there only exist decoder actions for all values of $y_1 \in [x_1^L(0), x_1^U(0)]$ and $y_2 = 0$, and we wish to obtain the best response of the encoder to these decoder actions. For any given decoder actions $\tilde{\boldsymbol{y}}$ and $\bar{\boldsymbol{y}}$ satisfying $\tilde{\boldsymbol{y}} \neq \bar{\boldsymbol{y}}$, it must be that $\tilde{y}_1 \neq \bar{y}_1$ as every decoder action is assumed to have $y_2 = 0$. Note also that for any given observation $\boldsymbol{x}$ in the support of the joint distribution, there exists a decoder action $\boldsymbol{y} = [y_1, y_2]^T$ with $y_1 = x_1$ and $y_2 = 0$. These imply that if a given observation $\boldsymbol{x}$ satisfies $x_1 = \tilde{y}_1$, it follows that

$$c_t^e(\boldsymbol{x}, \tilde{\boldsymbol{y}}) = (x_1 - \tilde{y}_1)^2 + (x_2 - \tilde{y}_2 - \tilde{b})^2$$
$$< (x_1 - \bar{y}_1)^2 + (x_2 - \bar{y}_2 - \tilde{b})^2 = c_t^e(\boldsymbol{x}, \bar{\boldsymbol{y}})$$

for any decoder actions $\bar{\boldsymbol{y}}$ and $\tilde{\boldsymbol{y}}$ satisfying $\bar{\boldsymbol{y}} \neq \tilde{\boldsymbol{y}}$. Therefore, if we denote the bin corresponding to a decoder action $\boldsymbol{y}$ by $B_{\boldsymbol{y}}$, we get $B_{\boldsymbol{y}} = \{\boldsymbol{x} \mid x_1 = y_1 \text{ and } x_2 \in \mathbb{R}\}$. This means that the best response of the encoder is to reveal the value of X_1 completely without giving any additional information. Hence, the encoding policy $z = \tilde{\gamma}^e(\boldsymbol{x}) = x_1$ and the decoding policy $\boldsymbol{y} = \tilde{\gamma}^d(z) = [z, 0]^T$ form a Nash equilibrium as they are best response maps of each other.

(iv) From Kagan et al. (1973, Theorem 5.3.1), we know that $\mathbb{E}[X_2|X_1 = x_1] = 0$ for all x_1 regardless of the source distribution when $b_1 = -b_2$. In this case, a similar analysis can be carried out to show that there always exists an informative linear Nash equilibrium for any source distribution when $b_1 = -b_2$.

Appendix G. Proof of Theorem 17

By applying a suitable linear transformation of variables, one can obtain an equivalent problem. In this equivalent problem, the aim is to convey a sequence of independent Gaussian sources, and the bias is zero for all sources except one. Due to the independence of these sources, the problem decouples. By using the result of Kazıklı et al. (2022, Theorem 4), it follows that there exists a Nash equilibrium where the encoder uses a quantization policy with any number of bins for the source corresponding to a non-zero bias. On the other hand, since the bias is zero for all other sources, revealing these sources leads to a Nash equilibrium.

While it is possible to apply a transformation of variables for any $n \geq 2$, we specify a transformation of variables for $n = 3$ dimensional scenario as an example. In particular, consider $\boldsymbol{X} = \mathcal{T}\boldsymbol{M}$ and $\boldsymbol{U} = \mathcal{T}^{-1}\boldsymbol{Y}$ where

$$\mathcal{T} \triangleq \frac{1}{\sqrt{b_1^2 + b_2^2 + b_3^2}}$$

$$
\times \begin{bmatrix}
\frac{b_2\sqrt{b_1^2+b_2^2+b_3^2}}{\sqrt{b_1^2+b_2^2}} & \frac{-b_1\sqrt{b_1^2+b_2^2+b_3^2}}{\sqrt{b_1^2+b_2^2}} & 0 \\[2ex]
\frac{b_1 b_3}{\sqrt{b_1^2+b_2^2}} & \frac{b_2 b_3}{\sqrt{b_1^2+b_2^2}} & -\sqrt{b_1^2+b_2^2} \\[2ex]
b_1 & b_2 & b_3
\end{bmatrix} = (\mathcal{T}^{-1})^T.
$$

Under this transformation of variables, the objective function of the encoder becomes

$$
\begin{aligned}
c^e(\boldsymbol{m}, \boldsymbol{u}) &= (x_1 - y_1)^2 + (x_2 - y_2)^2 \\
&\quad + (x_3 - y_3 - (b_1^2 + b_2^2 + b_3^2)^{1/2})^2 \triangleq c_t^e(\boldsymbol{x}, \boldsymbol{y}).
\end{aligned}
$$

For the objective function of the decoder, we get $c^d(\boldsymbol{m}, \boldsymbol{u}) = \|\boldsymbol{m} - \boldsymbol{u}\|^2 = \|\boldsymbol{x} - \boldsymbol{y}\|^2 \triangleq c_t^d(\boldsymbol{x}, \boldsymbol{y})$. Since X_1, X_2, X_3 are independent random variables, we obtain decoupled one-dimensional cheap talk problems where the biases for X_1 and X_2 are zero, and the bias for X_3 is non-zero. Thus, revealing X_1 and/or X_2 and applying a signaling game policy for X_3 yield a Nash equilibrium.

Appendix H. Proof of Theorem 18

If the condition in the statement of the theorem holds, the sources $X_1 = b_1 M_2 - b_2 M_1$ and $X_2 = b_1 M_1 + b_2 M_2$ in the equivalent problem become independent. Therefore, in the equivalent problem, we get decoupled one-dimensional cheap talk problems where there is a zero bias for X_1 and a non-zero bias for X_2. Thus, revealing X_1 yields an informative linear Nash equilibrium.

Appendix I. Proof of Theorem 19

If the encoder uses $M_1, \ldots, M_{\tilde{n}}$ with $\tilde{n} < n$ in constructing its linear policy as described below and gives no information related to $M_{\tilde{n}+1}, \ldots, M_n$, we obtain a linear Nash equilibrium where the encoder reveals $(\tilde{n} - 1)$ dimensions completely. Suppose without loss of generality that $\mathbb{E}[M_1] = 0$. We first apply a linear transformation of variables. In a similar manner to the two-dimensional case, we obtain an equivalent problem in a transformed coordinate system. In particular, let

$$
\mathcal{T} \triangleq \begin{bmatrix}
\frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} & 0 & \cdots & \\
\frac{1}{\sqrt{2\times 3}} & \frac{1}{\sqrt{2\times 3}} & \frac{-2}{\sqrt{2\times 3}} & 0 & \cdots \\
& & \vdots & & \\
\frac{1}{\sqrt{(n-1)\times n}} & \cdots & & \frac{1}{\sqrt{(n-1)\times n}} & \frac{-(n-1)}{\sqrt{(n-1)\times n}} \\
\frac{1}{\sqrt{n}} & \cdots & & & \frac{1}{\sqrt{n}}
\end{bmatrix}. \tag{35}
$$

Since $\mathcal{T}^{-1} = \mathcal{T}^T$ and $\mathcal{T}\boldsymbol{b} = [0, \ldots, 0, \sqrt{n}b]^T$, the cost function of the encoder for the equivalent problem becomes

$$
\begin{aligned}
c^e(\boldsymbol{m}, \boldsymbol{u}) &= \sum_{k=1}^n (m_k - u_k - b)^2 \\
&= \sum_{k=1}^{n-1} (x_k - y_k)^2 + (x_n - y_n - \sqrt{n}b)^2 \triangleq c_t^e(\boldsymbol{x}, \boldsymbol{y}).
\end{aligned}
$$

The cost function of the decoder in this transformed coordinate system is given by $c^d(\boldsymbol{m}, \boldsymbol{u}) = \|\boldsymbol{m} - \boldsymbol{u}\|^2 = \|\boldsymbol{x} - \boldsymbol{y}\|^2 \triangleq c_t^d(\boldsymbol{x}, \boldsymbol{y})$. It is seen that there is no bias for $X_1, \ldots, X_{n-1}$, and there is a non-zero bias for X_n considering the cost function of the encoder in this transformed coordinate system.

We note that revealing $X_1, \ldots, X_{n-1}$ is equivalent to revealing $M_1 - M_2, M_2 - M_3, \ldots, M_{n-1} - M_n$. In the following, we show that if the encoder reveals these random variables, the optimal estimate for X_n at the decoder becomes zero. Towards that goal, let $\tilde{M}_1 \triangleq -M_n, \tilde{M}_2 \triangleq -M_{n-1}, \ldots, \tilde{M}_n \triangleq -M_1$ and observe that

$\tilde{M}_1, \ldots, \tilde{M}_n$ have the same distribution as $M_1, \ldots, M_n$ due to the symmetry of the source distribution. Hence, we get

$$
\begin{aligned}
&\mathbb{E}[M_1 + \cdots + M_n | M_1 - M_2, \ldots, M_{n-1} - M_n] \\
&= -\mathbb{E}[\tilde{M}_n + \cdots + \tilde{M}_1 | \tilde{M}_{n-1} - \tilde{M}_n, \ldots, \tilde{M}_1 - \tilde{M}_2].
\end{aligned}
$$

This proves that the conditional mean of X_n given that $X_1, \ldots, X_{n-1}$ are revealed is zero. We know that in the transformed coordinate system, there is a non-zero bias only for X_n. If the encoder reveals $X_1, \ldots, X_{n-1}$ completely, the best response of the decoder is to use these revealed parameters as the corresponding estimates, and the action taken for X_n is zero. In other words, if we have $\boldsymbol{z} = \tilde{\gamma}^e(\boldsymbol{x}) = [x_1, \ldots, x_{n-1}]^T$ as the encoding policy, the best response of the decoder becomes $\boldsymbol{y} = \tilde{\gamma}^d(\boldsymbol{z}) = [z_1, \ldots, z_{n-1}, 0]^T$. Since we consider the Nash equilibrium concept, it is required to take the best response of the encoder into account, as well. Towards that goal, one can construct quantization bins as in the proof of Theorem 13 to prove the result. In fact, it is also possible to see the result by only looking at the policies. In particular, suppose that the decoder uses the policy $\boldsymbol{y} = \tilde{\gamma}^d(\boldsymbol{z}) = [z_1, \ldots, z_{n-1}, 0]^T$. This means that the decoder uses $y_1, \ldots, y_{n-1}$ as its estimates for $X_1, \ldots, X_{n-1}$, and the estimate for X_n is zero regardless of the transmitted message. Since the encoder wishes accurate estimations of $X_1, \ldots, X_{n-1}$ at the decoder without any bias, the best response of the encoder to the decoding policy $\boldsymbol{y} = \tilde{\gamma}^d(\boldsymbol{z}) = [z_1, \ldots, z_{n-1}, 0]^T$ becomes $\boldsymbol{z} = \tilde{\gamma}^e(\boldsymbol{x}) = [x_1, \ldots, x_{n-1}]^T$. Although the encoder has a bias regarding X_n, the encoder cannot affect the corresponding estimate at the decoder since the decoder action y_n is zero regardless of the encoded message. This implies that the encoding policy $\boldsymbol{z} = \tilde{\gamma}^e(\boldsymbol{x}) = [x_1, \ldots, x_{n-1}]^T$ and the decoding policy $\boldsymbol{y} = \tilde{\gamma}^d(\boldsymbol{z}) = [z_1, \ldots, z_{n-1}, 0]^T$ are best response maps of each other. Hence, an encoding policy that completely reveals $X_1, \ldots, X_{n-1}$ leads to a Nash equilibrium.

Appendix J. Proof of Theorem 21

Suppose without loss of generality that $\mathbb{E}[M_1] = 0$. Consider the transformation of variables in (35). We obtain an equivalent problem as in Lemma 9 where the linear transformation $\mathcal{T}$ is fixed as in (35), and the encoder and decoder design $\tilde{\gamma}^e(\cdot)$ and $\tilde{\gamma}^d(\cdot)$, respectively. The random variables $X_1, \ldots, X_n$ defined by (35) are i.i.d. and follow the same distribution as $M_1, \ldots, M_n$. Thus, the problem decouples to n one-dimensional cheap talk problems where there is a non-zero bias only for one of the problems. Suppose that the encoder does not reveal information related to X_n, which is the source corresponding to a non-zero bias. In this case, the rate of the original problem is identical to that of a team theoretic problem with $(n - 1)$ i.i.d. sources. Since the term that contributes to the objective of the encoder is $\mathbb{E}[(X_n - Y_n - \sqrt{n}b)^2] = nb^2$ in the case that no information is conveyed related to X_n, we obtain an additional b^2 term for the encoder's distortion bound. As a result, we get

$$
\frac{\mathbb{E}[\sum_{i=1}^n (M_i - U_i - b)^2]}{n} = \frac{\mathbb{E}[\sum_{i=1}^{n-1} (X_i - Y_i)^2]}{n} + b^2.
$$

By taking the limit of both sides, we have

$$
\begin{aligned}
&\lim_{n\to\infty} \frac{\mathbb{E}[\sum_{i=1}^n (M_i - U_i - b)^2]}{n} \\
&= \lim_{n\to\infty} \frac{\mathbb{E}[\sum_{i=1}^{n-1} (X_i - Y_i)^2]}{n} + b^2 \\
&= \lim_{n\to\infty} \frac{\mathbb{E}[\sum_{i=1}^{n-1} (X_i - Y_i)^2]}{n-1} + b^2,
\end{aligned}
$$

where the last equality follows since $n \to \infty$. Hence, we obtain a reduced team theoretic problem where the encoder wishes to convey an i.i.d. source $X_1, \ldots, X_{n-1}$ with a zero bias. Therefore, if

the pair (R_T, D_T) is achievable in a team theoretic setup, $R = R_T$ and $D_e \geq D_T + b^2$ are achievable for the original game theoretic setup. For the distortion bound of the decoder, we use the relation $\mathbb{E}[\sum_{i=1}^{n}(M_i - U_i - b)^2] = \mathbb{E}[\sum_{i=1}^{n}(M_i - U_i)^2] + b^2 n$.

Appendix K. Proof of Theorem 24

If we take $D_T = \min\{D_e - b^2, D_d\}$ for the team theoretic setup in Theorem 21, the corresponding rate R_T is achievable for the game theoretic setup. Thus, we get

$$\inf\{R \mid (R, D_e, D_d) \text{ is achievable}\}$$
$$\leq \inf\{R \mid (R, \min\{D_e - b^2, D_d\}) \text{ is achievable}\},$$

which leads to the bound in the statement of the theorem via Lemma 23.

References

Akyol, E., Langbort, C., & Başar, T. (2017). Information-theoretic approach to strategic communication as a hierarchical game. *Proceedings of the IEEE*, *105*(2), 205–218.

Ambrus, A., & Takahashi, S. (2008). Multi-sender cheap talk with restricted state spaces. *Theoretical Economics*, *3*(1), 1–27.

Başar, T. (1985). An equilibrium theory for multiperson decision making with multiple probabilistic models. *IEEE Transactions on Automatic Control*, *30*(2), 118–132.

Başar, T., & Olsder, G. J. (1999). *SIAM classics in applied mathematics, Dynamic noncooperative game theory*. Philadelphia, PA.

Bassan, B., Gossner, O., Scarsini, M., & Zamir, S. (2003). Positive value of information in games. *International Journal of Game Theory*, *32*, 17–31.

Battaglini, M. (2002). Multiple referrals and multidimensional cheap talk. *Econometrica*, *70*(4), 1379–1401.

Boyd, S., & Vandenberghe, L. (2004). *Convex optimization*. Cambridge, UK: Cambridge University Press.

Chakraborty, A., & Harbaugh, R. (2007). Comparative cheap talk. *Journal of Economic Theory*, *132*(1), 70–94.

Cover, T. M., & Thomas, J. A. (2006). *Elements of information theory*. New Jersey: John Wiley & Sons.

Crawford, V. P., & Sobel, J. (1982). Strategic information transmission. *Econometrica*, *50*(6), 1431–1451.

Harsanyi, J. C., & Selten, R. (1988). *A general theory of equilibrium selection in games*. Cambridge, Massachusets: MIT Press.

Hiriart-Urruty, J.-B. (1980). Mean value theorems in nonsmooth analysis. *Numerical Functional Analysis and Optimization*, *2*(1), 1–30.

Kagan, A. M., Linnik, Y. V., & Rao, C. R. (1973). *Characterization problems in mathematical statistics*. New York: John Wiley & Sons.

Kavitha, V., Altman, E., El-Azouzi, R., & Sundaresan, R. (2012). Opportunistic scheduling in cellular systems in the presence of noncooperative mobiles. *IEEE Transactions on Information Theory*, *58*(3), 1757–1773.

Kazıklı, E., Gezici, S., & Yüksel, S. (2021). Signaling games in higher dimensions: Geometric properties of equilibrium partitions. In *International symposium on modeling and optimization in mobile, ad hoc, and wireless networks (WiOpt)* (pp. 256–263).

Kazıklı, E., Sarıtaş, S., Gezici, S., Linder, T., & Yüksel, S. (2022). Signaling games for log-concave distributions: Number of bins and properties of equilibria. *IEEE Transactions on Information Theory*, *68*(3), 1731–1757.

Kazıklı, E., Sarıtaş, S., Gezici, S., & Yüksel, S. (2023). Quadratic signaling with prior mismatch at an encoder and decoder: Equilibria, continuity, and robustness properties. *IEEE Transactions on Automatic Control*, *68*(2), 705–720.

Larrousse, B., Beaude, O., & Lasaulce, S. (2014). Crawford-Sobel meet Lloyd-Max on the grid. In *IEEE international conference on acoustics, speech and signal processing (ICASSP)* (pp. 6127–6131).

Le Treust, M., & Tomala, T. (2019). Persuasion with limited communication capacity. *Journal of Economic Theory*, *184*, Article 104940.

Levy, G., & Razin, R. (2007). On the limits of communication in multidimensional cheap talk: A comment. *Econometrica*, *75*(3), 885–893.

Li, Z., Dán, G., & Liu, D. (2020). A game theoretic analysis of LQG control under adversarial attack. In *IEEE conference on decision and control (CDC)* (pp. 1632–1639).

Miura, S. (2014). Multidimensional cheap talk with sequential messages. *Games and Economic Behavior*, *87*, 419–441.

Sanjari, S., Başar, T., & Yüksel, S. (2021). Isomorphism properties of optimality and equilibrium solutions under equivalent information structure transformations II: Stochastic dynamic games. arXiv:2104.06406.

Sarıtaş, S., Dán, G., & Sandberg, H. (2020). Passive fault-tolerant estimation under strategic adversarial bias. In *American control conference (ACC)* (pp. 4644–4651).

Sarıtaş, S., Gezici, S., & Yüksel, S. (2019). Hypothesis testing under subjective priors and costs as a signaling game. *IEEE Transactions on Signal Processing*, *67*(19), 5169–5183.

Sarıtaş, S., Yüksel, S., & Gezici, S. (2017). Quadratic multi-dimensional signaling games and affine equilibria. *IEEE Transactions on Automatic Control*, *62*(2), 605–619.

Sarıtaş, S., Yüksel, S., & Gezici, S. (2020). Dynamic signaling games with quadratic criteria under Nash and Stackelberg equilibria. *Automatica*, *115*, Article 108883.

Sayın, M. O., Akyol, E., & Başar, T. (2019). Hierarchical multistage Gaussian signaling games in noncooperative communication and control systems. *Automatica*, *107*, 9–20.

Sémirat, S. (2019). Strategic information transmission despite conflict. *International Journal of Game Theory*, *48*, 921–956.

Sobel, J. (2020). Signaling games. In M. Sotomayor, D. Pérez-Castrillo, & F. Castiglione (Eds.), *Complex social and behavioral systems : Game theory and agent-based models* (pp. 251–268). New York, NY: Springer.

Teixeira, A., Shames, I., Sandberg, H., & Johansson, K. H. (2015). A secure control framework for resource-limited adversaries. *Automatica*, *51*, 135–148.

Vora, A. S., & Kulkarni, A. A. (2020a). Achievable rates for strategic communication. In *IEEE international symposium on information theory (ISIT)* (pp. 1379–1384).

Vora, A. S., & Kulkarni, A. A. (2020b). Information extraction from a strategic sender over a noisy channel. In *IEEE conference on decision and control (CDC)* (pp. 354–359).

Ertan Kazıklı received the B.Sc. degree in Electrical and Electronics Engineering from Bilkent University, Turkey, in 2012, the M.Sc. degree in Computer Science from École Polytechnique Fédérale de Lausanne (EPFL), Switzerland, in 2015, and the Ph.D. degree in Electrical and Electronics Engineering from Bilkent University, Turkey, in 2021. Between 2021–2022, he was a Post-doctoral Researcher at the Department of Mathematics and Statistics at Queen's University in Canada. In 2022, he joined the Electrical and Electronics Engineering Department at TOBB University of Economics and Technology, where he is currently working as an Assistant Professor. His research interests include game theory, detection and estimation theory, communication systems, and information theory.

Sinan Gezici received the B.Sc. degree in electrical and electronics engineering from Bilkent University, Ankara, Turkey, in 2001, and the Ph.D. degree in electrical engineering from Princeton University, Princeton, NJ, USA, in 2006. From 2006 to 2007, he worked at Mitsubishi Electric Research Laboratories, Cambridge, MA, USA. Since 2007, he has been with the Department of Electrical and Electronics Engineering at Bilkent University, where he is currently a Professor. Dr. Gezici's research interests are in the areas of detection and estimation theory, visible light systems, wireless communications, and localization systems. Among his publications in these areas is the book Ultra-wideband Positioning Systems: Theoretical Limits, Ranging Algorithms, and Protocols (Cambridge University Press, 2008). Dr. Gezici has been an associate editor for IEEE Transactions on Vehicular Technology, IEEE Transactions on Communications, IEEE Wireless Communications Letters, and Journal of Communications and Networks.

Serdar Yüksel received his B.Sc. degree in Electrical and Electronics Engineering from Bilkent University; M.S. and Ph.D. degrees in Electrical and Computer Engineering from the University of Illinois at Urbana-Champaign in 2003 and 2006, respectively. He was a post-doctoral researcher at Yale University before joining the Department of Mathematics and Statistics at Queen's University. His research interests are on stochastic control, decentralized control, information theory, and probability. He has been an Associate Editor for the IEEE Transactions on Automatic Control, Automatica, Systems and Control Letters, and Mathematics of Control, Signals and Systems.