

Signaling Games in Higher Dimensions: Geometric Properties of Equilibrium Partitions

Ertan Kazıklı

*Electrical and Electronics Engineering
Bilkent University
Ankara, Turkey
kazikli@ee.bilkent.edu.tr*

Sinan Gezici

*Electrical and Electronics Engineering
Bilkent University
Ankara, Turkey
gezici@ee.bilkent.edu.tr*

Serdar Yüksel

*Mathematics and Statistics
Queen's University
Kingston, Ontario, Canada
yuksel@mast.queensu.ca*

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. We investigate a signaling game problem where an encoder observes a multi-dimensional source and conveys a message to a decoder, and the quadratic objectives of the encoder and decoder are misaligned due to a bias vector. For the scalar case, Crawford and Sobel in their seminal paper, show that under certain technical assumptions an encoding policy must be a quantization policy at any Nash equilibrium. We first provide a set of geometry conditions that needs to be satisfied in equilibrium considering any multi-dimensional source. Then, we consider multi-dimensional sources with independent and identically distributed components and completely characterize conditions under which a Nash equilibrium with a linear encoder exists. In particular, we show that if the components of the bias vector are not equal in magnitude, then there exists a linear equilibrium if and only if the source distribution is Gaussian. On the other hand, for a linear equilibrium to exist in the case of equal bias components, it is required that the source density is symmetric around its mean. 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 their team theoretic counterpart.

Index Terms—Signaling games, multi-dimensional cheap talk, Nash equilibrium, rate-distortion theory.

I. INTRODUCTION AND SYSTEM MODEL

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 cannot 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, there exist setups where the decision makers do not share a common goal [1]–[13]. In these studies, there are two main themes which lead to misaligned objectives for the decision makers. In the first theme, the decision makers employ different costs function as their goals are inconsistent, e.g., a decision maker wishes to mislead another decision maker [1]–[11]. The second theme is concerned with the case when the decision makers have subjective beliefs on probability distributions of

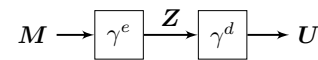


Fig. 1. Communication setting.

unknown parameters, which leads to misaligned objectives for the decision makers even though they employ the same cost function [6], [12], [13]. These both lead to a game theoretic setup where a suitable equilibrium concept such as the Nash equilibrium and Stackelberg equilibrium is to be used to analyze the system. In this context, Crawford and Sobel in their seminal paper [14] introduce a cheap talk problem where a better informed encoder communicates with a decoder, and the encoder and decoder have misaligned objectives due to a bias term. Here, under certain technical conditions, although there is no restriction imposed by the system, the encoder must employ quantization policies at any Nash equilibrium. Thus, it is not possible to have an equilibrium with a fully revealing encoder whereas there may exist a partition equilibrium with a quantized encoder. 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. To put it differently, if the encoder and decoder wish to operate at a Nash equilibrium, then the encoder cannot reveal its private information completely to the decoder.

In this paper, our aim is to investigate a multi-dimensional extension of Crawford and Sobel's formulation where the encoder jointly encodes its multi-dimensional observation to a message observed by the decoder. For the multi-dimensional source setting, an important problem is to see whether there exists a linear Nash equilibrium. Accordingly, we investigate whether there exists an equilibrium with a linear encoder for the multi-dimensional cheap talk problem with sources consisting of independent and identically distributed (i.i.d.) random variables. In contrast with the one-dimensional scenario, having more than one dimension allows the encoder to employ a linear encoder under certain conditions and we completely characterize these conditions considering i.i.d. sources.

In this paper, we consider the following multi-dimensional

cheap talk problem where an encoder and a decoder communicates. The encoder knows the value of an n -dimensional random vector $\mathbf{M} = [M_1, \dots, M_n]^T$ where M_1, \dots, M_n are \mathbb{M} -valued i.i.d. random variables.¹ The encoder conveys a message $\mathbf{Z} = [Z_1, \dots, Z_n]^T$ via an encoding policy $\gamma^e(\cdot)$, i.e., $\mathbf{Z} = \gamma^e(\mathbf{M})$, where Z_i is a \mathbb{Z}_i -valued random variable for $i = 1, \dots, n$. The decoder directly observes \mathbf{Z} and takes an action $\mathbf{U} = [U_1, \dots, U_n]^T$ where U_i 's are \mathbb{M} -valued random variables. The aim of the encoder is to minimize $J^e(\gamma^e, \gamma^d) = \mathbb{E}[c^e(\mathbf{M}, \mathbf{U})]$ where

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

In (1), \mathbf{b} denotes a deterministic bias vector which is common knowledge between the players and quantifies the degree of misalignment between the objective functions of the encoder and decoder. On the other hand, the decoder wishes to estimate the source random vector as accurately as possible, and thus, its objective function does not include a bias vector. In particular, the aim of the decoder is to minimize $J^d(\gamma^e, \gamma^d) = \mathbb{E}[c^d(\mathbf{M}, \mathbf{U})]$ where

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

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 if [15]

$$\begin{aligned} J^e(\gamma^{*,e}, \gamma^{*,d}) &\leq J^e(\gamma^e, \gamma^{*,d}) \text{ for all } \gamma^e \in \Gamma^e, \\ J^d(\gamma^{*,e}, \gamma^{*,d}) &\leq J^d(\gamma^{*,e}, \gamma^d) \text{ for all } \gamma^d \in \Gamma^d, \end{aligned} \quad (3)$$

where Γ^e and Γ^d are the sets of all deterministic (and Borel measurable) functions from \mathbb{M}^n to $(\mathbb{Z}_1, \dots, \mathbb{Z}_n)$ and from $(\mathbb{Z}_1, \dots, \mathbb{Z}_n)$ to \mathbb{M}^n , respectively.

We make the following assumption for the source distribution.

Assumption 1. Considering each component M_i of the source random vector \mathbf{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 $\mathbb{E}[\mathbf{M} | \mathbf{M} \in C]$ must be in the interior of set C . We use this implication while proving some of the results in the paper.

An important implication of our results is related to information theoretic limits of the cheap talk problem. This problem can be formulated as the maximum rate of communication

between an encoder and decoder in a multi-dimensional signaling setup where the rate is defined as the number of distinct messages per source symbol. This is an important problem in communication theory where the objectives are aligned but there may be an upper bound on the achievable communication rate due to system requirements such as a power constraint at the encoder and having a noisy channel. In a communication theoretic setup with a channel, Shannon's channel capacity defines the maximum rate of communication and forms the foundations of many results in information theory [16]. In particular, one wishes to design a communication system achieving this channel capacity in order to convey as much information as possible. On the other hand, an interesting question arises in a game theoretic setup: Does there exist an upper bound on achievable rate of communication in an equilibrium with an encoder and decoder where the encoder and decoder have misaligned objectives? In certain cases, our analysis gives a conclusive answer to this question for the multi-dimensional cheap talk setup. In particular, if there exists an equilibrium with a linear encoder, then there is no upper bound on the maximum rate of communication. Note that we consider the Nash setup for such an information theoretic problem. We refer the reader to [8] for a Stackelberg setup.

II. LITERATURE REVIEW

Investigation of communication between strategic agents which have misaligned objectives has gained a significant attention in recent control and communication theory literature [1]–[7], [9]–[12] in addition to a significant amount of work in the economics literature (see [17] for a literature survey). Research in communication between strategic agents dates back to the seminal work of Crawford and Sobel in the economics literature [14] which considers the Nash equilibrium concept. In contrast to simultaneous-move nature of the Nash equilibrium where each player announces their strategies at the same time, one can also consider a sequential-move setup by analyzing the Stackelberg equilibrium where there is an order in the decision making procedure. An important line of work in this context is the Bayesian persuasion problem initiated by Kamenica and Gentzkow [18] where a sender makes and announces its decision first, and commits to this information revelation policy. An important distinction between the Nash setup and the Stackelberg setup is the absence of commitment assumption in the Nash setup. Thus, while determining whether to use Nash or Stackelberg equilibria analysis, one should see if this commitment assumption is reasonable or not for the problem at hand.

Signaling game problems are investigated in various papers in the literature [1]–[12]. The work in [2] investigates signaling game setups under Nash and Stackelberg equilibria concepts where a sender with a deterministic bias term communicates with a receiver. A specific result in [2, Theorem 3.4] is related to our work where it is shown that for the multi-dimensional cheap talk problem a Nash equilibrium can be a partition (quantized) equilibrium or a linear equilibrium by providing specific examples. The work in [1] considers a problem under

¹While we present general results for n -dimensional cheap talk problem for any $n \geq 2$, some of our results focus on the case of two-dimensional cheap talk where $n = 2$. Nevertheless, it is also possible to generalize these results to cheap talk with an arbitrary number of i.i.d. sources by using the results for the two-dimensional scenario.

the Stackelberg equilibrium concept where the bias term at the sender is modeled as a random variable. In [11], for the one-dimensional cheap talk problem, properties related to number of bins at the equilibrium are analyzed for certain classes of source distributions such as a log-concave source and a source satisfying certain assumptions at the tails of the distribution. An important result from [11] is that for a source distribution whose support is \mathbb{R} and satisfying certain assumptions, there does not exist an upper bound on the number of bins in equilibrium, and that while equilibria are to be quantized more bins lead to smaller expected costs both for the encoder and the decoder. Another related work in [5] investigates multi-stage cheap talk problems under Nash and Stackelberg equilibria where the objectives involve mean squared error with a deterministic bias at the encoder. As an application, [19] applies Crawford and Sobel's formulation to a smart grid problem by modeling the interaction between a consumer and an electricity aggregator as a signaling game. As another application, strategic users in a cellular network may wish to misreport their channel conditions to the base station for their own benefit and the presence of strategic decision makers leads to a signaling game problem [20].

A set of related papers investigates information theoretic limits of communication under a Stackelberg setup where there is a commitment assumption for one of the players [7]–[10]. In particular, the works [7], [8] focus on the case where the sender acts first and commits to its policy as in the Bayesian persuasion setup. The authors provide an information theoretic perspective on the Bayesian persuasion problem by considering a communication setup where the encoder makes a sequence of observations from an i.i.d. information source. On the other hand, [9], [10] consider the case when the receiver has a commitment assumption and derive the information extraction capacity for such a setup. This type of commitment assumption has applications in designing optimal questionnaires by which information is collected from strategic agents [21].

Multidimensional cheap talk problems have also been considered in the economics literature [22]–[24]. For instance, [22] investigates a two-dimensional scenario where an encoder communicates with a decoder. Different from our work, in [22], 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 decoder action in the second dimension does not matter. In this case, [22] shows the existence of an upper bound on the number of decoder actions. In addition, the work in [23] considers a multidimensional cheap talk problem when there are two senders referred to as experts and a receiver referred to as a policy-maker. These decision makers including the policy-maker have inconsistent ideal points regarding an outcome of the game and these ideal point vectors can be viewed as bias vectors as in our framework. They show that if these ideal point vectors of the experts in the multidimensional space are linearly independent, then there exist fully revealing equilibria. While [23] studies

conditions on the existence of fully revealing equilibria, our focus instead is on the characterization of equilibrium partitions in general; as in the case with a single encoder/sender we do not have full revelation in general. More specifically, we focus on a scenario with a single encoder which jointly encodes its multi-dimensional observation and employs a single quadratic cost function.

III. NASH EQUILIBRIA

A. Necessary Conditions for Equilibria

We first present some preliminary results which are useful for providing conditions that ensure the existence of an equilibrium with a linear encoding policy. To begin with, due to the definition of $J^d(\gamma^e, \gamma^d)$, we know that for a given encoding policy $z = \gamma^e(\mathbf{m})$, the decoder takes the action $\mathbf{u} = \mathbb{E}[\mathbf{M} | \gamma^e(\mathbf{M}) = z]$. All such possible values of \mathbf{u} at the equilibria are called the set of decoder actions. While investigating the existence of a linear equilibrium, we make a frequent use of the following geometric definition regarding the set of decoder actions in equilibrium.

Definition 1. We say that a non-empty set of decoder actions forms a *continuum* if it is a closed and connected set (i.e., it cannot be expressed as the union of two or more disjoint closed sets.)

The following lemma derives a geometry condition that any two decoder actions at a Nash equilibrium need to satisfy. In addition to its usefulness while deriving conditions for the existence of linear equilibria, this result is important on its own as it provides a necessary condition for a Nash equilibrium in terms of the induced decoder actions. 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 1. Consider n -dimensional cheap talk problem with the source random vector $\mathbf{M} = [M_1, \dots, M_n]^T$ where each element of \mathbf{M} can have different distributions and can be dependent or independent. Let \mathcal{B}^α and \mathcal{B}^β be two bins and let $\mathbf{u}^\alpha = \mathbb{E}[\mathbf{M} | \mathbf{M} \in \mathcal{B}^\alpha]$ and $\mathbf{u}^\beta = \mathbb{E}[\mathbf{M} | \mathbf{M} \in \mathcal{B}^\beta]$ denote their centroids which are the decoder actions taken when the encoder reveals $\mathbf{M} \in \mathcal{B}^\alpha$ and $\mathbf{M} \in \mathcal{B}^\beta$, respectively.

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

$$2 \|(\mathbf{u}^\beta - \mathbf{u}^\alpha)^T \mathbf{b}\| \leq \|\mathbf{u}^\beta - \mathbf{u}^\alpha\|^2. \quad (4)$$

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

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

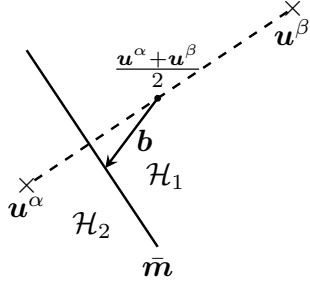


Fig. 2. Illustration of half spaces induced by decoder actions \mathbf{u}^α and \mathbf{u}^β in Lemma 1 considering a two-dimensional scenario. Here, the crosses represent the decoder actions and the arrow represents the bias vector. These decoder actions and the bias vector lead to a line of $\bar{\mathbf{m}}$ values for which the encoder is indifferent between reporting these observations as \mathbf{u}^α or \mathbf{u}^β . Here, the half spaces \mathcal{H}_1 and \mathcal{H}_2 cannot intersect with \mathcal{B}^α and \mathcal{B}^β , respectively.

and $h(\mathbf{m}) = 0$ defines the hyperplane on which the encoder is indifferent between either decoder actions, i.e., these \mathbf{m} values can belong to both \mathcal{B}^α and \mathcal{B}^β .

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

Proof Sketch. We provide a brief explanation here. For the complete proof, see [25]. Fig. 2 illustrates the open half spaces induced by decoder actions \mathbf{u}^α and \mathbf{u}^β . The encoder wishes to report the observations in \mathcal{H}_1 as \mathbf{u}^β rather than as \mathbf{u}^α . Similarly, the encoder wishes to report the observations in \mathcal{H}_2 as \mathbf{u}^β rather than as \mathbf{u}^α . Due to the centroid condition, \mathcal{H}_2 cannot include \mathbf{u}^β , and \mathcal{H}_1 cannot include \mathbf{u}^α . This implies that the line of $\bar{\mathbf{m}}$ values for which the encoder is indifferent between reporting these observations as either of the decoder actions must be in between \mathbf{u}^α and \mathbf{u}^β . Then, by expressing this condition, one can obtain (4). The convexity of the quantization bins follows from the fact that \mathcal{H}_1 and \mathcal{H}_2 for any two decoder actions are half spaces that decompose the complete observation space. \square

Lemma 1 presents a geometry condition that any two decoder actions in equilibrium must satisfy. It is important to emphasize that this condition applies to any joint distribution for the multi-dimensional observations. In particular, the result of Lemma 1 holds even for joint distributions which are not independent and identically distributed since the derivation of the condition in Lemma 1 uses only the cost function of the encoder and does not use the joint distribution of the sources.

Now, we focus on equilibria with continuously connected set of decoder actions. We provide a necessary condition that an equilibrium with decoder actions supported on a continuum has to satisfy. An equilibrium with a continuum of decoder actions means that the encoding policy is not a quantization policy.

In the remainder of this subsection, we specialize to the case of two-dimensional cheap talk problem where $n = 2$.

Lemma 1 implies that for decoder actions \mathbf{u}^α and \mathbf{u}^β satisfying $(\mathbf{u}^\beta - \mathbf{u}^\alpha)^T \mathbf{b} = 0$, it is possible to make their distance $\|\mathbf{u}^\alpha - \mathbf{u}^\beta\|$ arbitrarily small. On the other hand, for

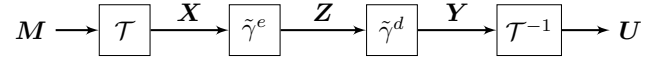


Fig. 3. Equivalent formulation where \mathcal{T} denotes the linear transformation specified by $X_1 = b_1 M_2 - b_2 M_1$ and $X_2 = b_1 M_1 + b_2 M_2$, and \mathcal{T}^{-1} denotes its inverse.

decoder actions \mathbf{u}^α and \mathbf{u}^β with $(\mathbf{u}^\beta - \mathbf{u}^\alpha)^T \mathbf{b} \neq 0$, since the distance $\|\mathbf{u}^\alpha - \mathbf{u}^\beta\|$ is lower bounded by a positive value, these decoder actions \mathbf{u}^α and \mathbf{u}^β cannot get arbitrarily close. This motivates an equivalent formulation by introducing the following transformation of variables. In particular, we define

$$\mathbf{X} = \mathcal{T}\mathbf{M}, \quad (6)$$

$$\mathbf{U} = \mathcal{T}^{-1}\mathbf{Y}, \quad (7)$$

where

$$\mathcal{T} = \begin{bmatrix} -b_2 & b_1 \\ b_1 & b_2 \end{bmatrix}, \quad \mathcal{T}^{-1} = \frac{1}{b_1^2 + b_2^2} \begin{bmatrix} -b_2 & b_1 \\ b_1 & b_2 \end{bmatrix}, \quad (8)$$

and $\mathbf{X} \triangleq [X_1, X_2]^T$ and $\mathbf{Y} \triangleq [Y_1, Y_2]^T$ respectively denote the observation at the encoder and the decoder action in the transformed coordinate system. The proposed equivalent formulation is depicted in Fig. 3 where the linear transformation \mathcal{T} and its inverse \mathcal{T}^{-1} is fixed, and the encoder and decoder simply design $\tilde{\gamma}^e(\cdot)$ and $\tilde{\gamma}^d(\cdot)$. In this coordinate system, the cost function of the encoder becomes

$$\begin{aligned} c^e(\mathbf{m}, \mathbf{u}) &= (\mathbf{m} - \mathbf{u} - \mathbf{b})^T (\mathbf{m} - \mathbf{u} - \mathbf{b}) \\ &= (\mathcal{T}^{-1} \mathcal{T}(\mathbf{m} - \mathbf{u} - \mathbf{b}))^T (\mathcal{T}^{-1} \mathcal{T}(\mathbf{m} - \mathbf{u} - \mathbf{b})) \\ &= (\mathcal{T}(\mathbf{m} - \mathbf{u} - \mathbf{b}))^T (\mathcal{T}^{-1})^T (\mathcal{T}^{-1}) (\mathcal{T}(\mathbf{m} - \mathbf{u} - \mathbf{b})). \end{aligned}$$

Since $(\mathcal{T}^{-1})^T (\mathcal{T}^{-1}) = (1/(b_1^2 + b_2^2))I$ with I denoting identity matrix, $\mathcal{T}\mathbf{b} = [0, b_1^2 + b_2^2]^T$, $\mathcal{T}\mathbf{m} = \mathbf{x}$ and $\mathcal{T}\mathbf{u} = \mathbf{y}$, we get

$$c^e(\mathbf{m}, \mathbf{u}) = \frac{(x_1 - y_1)^2}{\tilde{b}} + \frac{(x_2 - y_2 - \tilde{b})^2}{\tilde{b}} \triangleq \frac{c_t^e(\mathbf{x}, \mathbf{y})}{\tilde{b}}, \quad (9)$$

where

$$\tilde{b} \triangleq b_1^2 + b_2^2, \quad (10)$$

and the common factor of $(1/\tilde{b})$ is canceled in the definition of $c_t^e(\mathbf{x}, \mathbf{y})$ for notational convenience. This implies that there is no bias for the first coordinate whereas there is a bias of \tilde{b} for the second coordinate. In a similar manner, the cost function of the decoder in this coordinate system can be expressed as

$$\begin{aligned} c^d(\mathbf{m}, \mathbf{u}) &= (m_1 - u_1)^2 + (m_2 - u_2)^2 \\ &= \frac{(x_1 - y_1)^2}{\tilde{b}} + \frac{(x_2 - y_2)^2}{\tilde{b}} \triangleq \frac{c_t^d(\mathbf{x}, \mathbf{y})}{\tilde{b}}. \end{aligned} \quad (11)$$

With this transformation of variables, the aim of the encoder and decoder is to minimize $J^e(\tilde{\gamma}^e, \tilde{\gamma}^d) = \mathbb{E}[c_t^e(\mathbf{X}, \mathbf{Y})]$ and $J^d(\tilde{\gamma}^e, \tilde{\gamma}^d) = \mathbb{E}[c_t^d(\mathbf{X}, \mathbf{Y})]$, respectively. The expression in (11) shows that the cost of the decoder still involves a sum of squared error terms in the transformed coordinate system. This observation implies that for a given encoding policy, the best response of the decoder is still given by the conditional

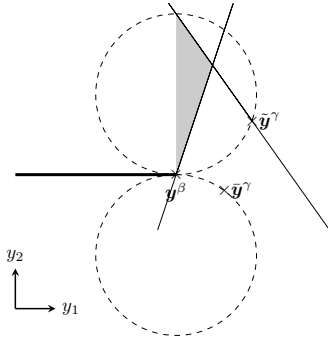


Fig. 4. Illustration of the proof technique employed in Lemma 4. Here, the horizontal solid line represents a continuum of decoder actions and the decoder action on this continuum with the largest y_1 coordinate is denoted by \mathbf{y}^β . Inside the dashed circles, there cannot be a decoder action due to the condition imposed by Lemma 1. If we place decoder actions $\tilde{\mathbf{y}}^\gamma$ and $\tilde{\mathbf{y}}^\gamma$ on the dashed circles, then the shaded area must be the bin for \mathbf{y}^β , i.e., \mathcal{B}^β . However, the centroid of this shaded area cannot be \mathbf{y}^β , which is a contradiction.

mean expressions in the transformed coordinate system. We summarize this result in the following lemma.

Lemma 2. For a fixed encoding function $\tilde{\gamma}^e(\mathbf{x})$, the optimal $\tilde{\gamma}^d(\cdot)$ that minimizes $J^d(\tilde{\gamma}^e, \tilde{\gamma}^d)$ is given by $\mathbb{E}[\mathbf{X}|\mathbf{Z} = \mathbf{z}]$.

Equipped with this transformation of variables and the cost functions in this transformed coordinate system, we are now ready to present our results on necessary conditions which need to be satisfied for any equilibrium with a continuum of decoder actions.

Lemma 3. At an equilibrium, any continuum of decoder actions must have a constant y_2 coordinate, i.e., a continuum of decoder actions can only be supported on $y_2 = \kappa$ for some κ in the support of X_2 .

Proof. See [25]. \square

Lemma 3 implies that a continuum of actions is allowed only in a specific direction that depends on the bias terms in the original coordinate system. There exists another constraint for such continuum of decoder actions as the next lemma shows.

Lemma 4. Let there be a continuum of decoder actions with a constant second coordinate, 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 $X_2 = \kappa$.

Proof Sketch. We briefly explain the proof technique here. For the complete proof, see [25]. Suppose that there exists a continuum of actions in equilibrium as illustrated in Fig. 4. Assume, by contradiction, that this continuum of actions have a discontinuity at \mathbf{y}^β . Let \mathcal{B}^β denote the bin corresponding to \mathbf{y}^β , i.e., $\mathbf{y}^\beta = \mathbb{E}[\mathbf{X}|\mathbf{X} \in \mathcal{B}^\beta]$. If we take the necessary condition in (4) into account, then there cannot exist any decoder actions inside the dashed circles in Fig. 4 in an equilibrium. Then, considering all possible decoder actions in

an equilibrium on the two-dimensional space, it is possible find a non-empty region of observations with $y_1^\beta < x_1$ and $\mathbf{x} \in \mathcal{B}^\beta$. In other words, in this non-empty region, the encoder always reports these observations as $\mathbf{X} \in \mathcal{B}^\beta$ since the decoder action \mathbf{y}^β induces the smallest cost for the encoder for these observations among all possible decoder actions in an equilibrium. Moreover, due to the decoder actions on the continuum, there cannot exist any observations with $x_1 < y_1^\beta$ and $\mathbf{x} \in \mathcal{B}^\beta$. This is a contradiction to $\mathbf{y}^\beta = \mathbb{E}[\mathbf{X}|\mathbf{X} \in \mathcal{B}^\beta]$ due to Assumption 1. \square

Lemma 3 states that a continuum of decoder actions must have a constant y_2 coordinate and Lemma 4 states that these 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 an equilibrium, as investigated in the next subsection.

B. Main Results

In this subsection, we present our main results on the existence of linear equilibria. In order to establish our main results, we employ Lemma 3 and Lemma 4 together with an interesting result from the literature known as Kagan-Linnik-Rao Theorem [26, Theorem 5.3.1]. This theorem characterizes conditions under which $\mathbb{E}[M_1 - \alpha M_2 | M_1 + \beta M_2]$ is a constant for i.i.d. M_1 and M_2 where α and β are non-zero constants.

Theorem 1. Consider 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 $b_1 = 0$ and/or $b_2 = 0$, there always exists an informative equilibrium with a linear encoder where the encoder completely reveals the source(s) corresponding to zero bias term(s).
- (ii) For $|b_1| \neq |b_2|$, there exists an informative equilibrium with a linear encoder if and only if the source distribution is Gaussian.
- (iii) For $b_1 = b_2$, there exists an informative 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 $b_1 = -b_2$, there always exists an informative equilibrium with a linear encoder regardless of the source distribution.

Proof Sketch. In the case of zero bias term(s), it is easy to verify that revealing the corresponding source(s) leads to a Nash equilibrium. This is a consequence of the fact that M_1 and M_2 are independent, and thus, the problem decouples into two one-dimensional cheap talk problems where one or both of the problems involve an encoder with a zero bias. In this case, an equilibrium may be attained without joint encoding of the sources due to the decoupled nature of the problem. In

the following, we treat the remaining cases where $b_1 \neq 0$ and $b_2 \neq 0$.

As mentioned earlier, main ingredients in the proof are Lemma 3 and Lemma 4 of this paper and Kagan-Linnik-Rao Theorem. From Lemma 3 and Lemma 4, we know that a continuum of actions must have a constant y_2 coordinate and must be supported for all values of y_1 in the support of X_1 given that $X_2 = y_2$. To conclude the results, we use conditions when $\mathbb{E}[X_2|X_1 = x_1] = \mathbb{E}[b_1 M_1 + b_2 M_2 | b_1 M_2 - b_2 M_1 = x_1]$ is a constant from [26, Theorem 5.3.1]. In particular, in the case when $|b_1| \neq |b_2|$, it is required to have a Gaussian distribution so that $\mathbb{E}[X_2|X_1 = x_1]$ is a constant. This implies that if $|b_1| \neq |b_2|$ and the source distribution is not Gaussian, then there does not exist a linear equilibrium. On the other hand, if the source distribution is Gaussian, one can employ independence of X_1 and X_2 to show the existence of a linear equilibrium in this case. In the case of $b_1 = b_2$, [26, Theorem 5.3.1] states that $\mathbb{E}[X_2|X_1 = x_1]$ is constant if and only if the source distribution is symmetric. Moreover, in the case of $b_1 = -b_2$, [26, Theorem 5.3.1] implies that $\mathbb{E}[X_2|X_1 = x_1]$ is constant regardless of the source distribution. In these cases, if the encoder reveals X_1 completely, then the best response of the decoder yields a single continuum of decoder actions. The best response of the encoder to this single continuum of decoder actions is given by revealing X_1 completely as otherwise the performance of the encoder degrades. Hence, it follows that revealing X_1 completely leads to a Nash equilibrium when $b_1 = b_2$ and the source distribution is symmetric, or when $b_1 = -b_2$. \square

Remark 1. Theorem 1 shows that depending on certain conditions there exists an informative 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, when the source distribution has bounded support, there exists an upper bound on the number of bins in equilibrium. 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 an equilibrium in the case of one-dimensional scenario can be non-informative, in the case of two-dimensional scenario with the same bias as in the one-dimensional scenario in both dimensions, it is possible to obtain an informative equilibrium when the source distribution is symmetric.

Remark 2. In the case of a Gaussian source, the problem decouples into two one-dimensional cheap talk problems. In particular, $X_1 = b_1 M_2 - b_2 M_1$ and $X_2 = b_1 M_1 + b_2 M_2$ become independent random variables when M_1 and M_2 are i.i.d. Gaussian. In fact, due to Darmois-Skitovich Theorem [26, 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 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 \hat{b} . From [11, 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 equilibrium with N bins. Thus, for two-dimensional cheap talk problem with a Gaussian source, there exists an equilibrium where the encoder reveals X_1 completely and applies a signaling game 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 [11, Theorem 4] and independence of the random variables in the transformed coordinate system, it follows that there exists an equilibrium where the encoder applies a quantization policy to this remaining random variable with any number of bins. Thus, we obtain the following result.

Theorem 2. *If the source distribution is n -dimensional i.i.d. Gaussian, then there exists an equilibrium with a linear encoding policy where the encoder reveals all or a subset of $(n - 1)$ dimensions completely in a transformed coordinate system, and applies a signaling game policy for the remaining dimension with any number of bins.*

Proof. See [25]. \square

IV. 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. In the special case that 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 equilibrium where the sender completely reveals $(n - 1)$ dimensions in a transformed coordinate system. For the remaining dimension, the encoder has to employ a quantization 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 a $(n - 1)$ -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 1. Consider the n -dimensional cheap talk problem with i.i.d. sources and $b = b_1 = \dots = b_n$. 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 $\gamma_n^e : \mathbb{M}^n \rightarrow \mathcal{Z}(n)$ with $|\mathcal{Z}(n)| \leq 2^{Rn}$ where $|\mathcal{Z}(n)|$ denotes the cardinality of $\mathcal{Z}(n)$.
- (ii) The decoder is given by $\gamma_n^d : \mathcal{Z}(n) \rightarrow \mathbb{M}^n$ such that

$$\lim_{n \rightarrow \infty} \frac{\mathbb{E} \left[\sum_{i=1}^n (M_i - U_i - b)^2 \right]}{n} \leq D_e, \quad (12)$$

$$\lim_{n \rightarrow \infty} \frac{\mathbb{E} \left[\sum_{i=1}^n (M_i - U_i)^2 \right]}{n} \leq D_d. \quad (13)$$

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 1 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 bias only for a single random variable. The following theorem uses this idea to relate achievable rate and distortion values of the original problem to that of a team theoretic problem. Before presenting the theorem, 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 $D_e \geq D_d + b^2$.

Theorem 3. Consider 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:

$$R = R_T, \quad (14)$$

$$D_e \geq D_T + b^2, \quad (15)$$

$$D_d \geq D_T. \quad (16)$$

Proof. See [25]. \square

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

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

By using the result of Theorem 3, we can upper bound the rate-distortion function for the Gaussian case, as stated in the following.

Theorem 4. Consider 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 is upper bounded by the following:

$$R(D_e, D_d) \leq \frac{1}{2} \log_2 \frac{\sigma^2}{\min\{D_d, D_e - b^2\}} \quad (18)$$

if $0 \leq \min\{D_d, D_e - b^2\} \leq \sigma^2$. On the other hand, we have $R(D_e, D_d) = 0$ if $\min\{D_d, D_e - b^2\} > \sigma^2$.

Proof. See [25]. \square

Remark 3. A related result can be found in [11, Theorem 8]. Here, it is shown that for one-dimensional cheap talk having more bins in the quantized encoding policy leads to reduced distortion values if the source distribution is log-concave. In other words, a large rate leads to better expected costs for both of the players under a log-concave source assumption, which holds for a Gaussian source.

Remark 4. For multi-dimensional i.i.d. Gaussian sources, we know that there exists an 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 [11, 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 an equilibrium with infinitely many bins applied to X_n due to [11, Theorem 13]. This implies that an equilibrium where X_1, \dots, X_{n-1} are revealed and a quantization policy with infinitely many bins applied to X_n corresponds to a payoff dominant Nash equilibrium [27].

V. CONCLUSION

We have analyzed a quadratic multi-dimensional cheap talk problem. First, we have derived general necessary 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 geometry condition that essentially prevents any two decoder actions being arbitrarily close to each other depending on their difference as vectors and the bias vector. Then, we have derived conditions under which a linear equilibrium exists considering i.i.d. sources. In particular, we have proven that if the bias vector has components which are not equal in magnitude, then there exists a linear equilibrium if and only if the source distribution is Gaussian. On the other hand, if the bias vector has components equal to each other, then there exists a linear equilibrium if and only if the source distribution is symmetric. In contrast, if the

components are equal in magnitude with different signs, then there always exists a linear equilibrium regardless of the source distribution. This complete characterization of conditions that ensure existence of linear equilibria has implications on the information theoretic limits of the cheap talk problem. If there exists a linear equilibrium, then there is no upper bound on the maximum rate of communication arising from the biased nature of the encoder. As future work, we plan to investigate information theoretic limits in the remaining cases when there does not exist a linear equilibrium.

REFERENCES

- [1] E. Akyol, C. Langbort, and T. Başar, "Information-theoretic approach to strategic communication as a hierarchical game," *Proceedings of the IEEE*, vol. 105, no. 2, pp. 205–218, 2017.
- [2] S. Saritaş, S. Yüksel, and S. Gezici, "Quadratic multi-dimensional signaling games and affine equilibria," *IEEE Transactions on Automatic Control*, vol. 62, no. 2, pp. 605–619, 2017.
- [3] F. Farokhi, A. M. H. Teixeira, and C. Langbort, "Estimation with strategic sensors," *IEEE Transactions on Automatic Control*, vol. 62, no. 2, pp. 724–739, 2017.
- [4] M. O. Sayin, E. Akyol, and T. Başar, "Hierarchical multistage Gaussian signaling games in noncooperative communication and control systems," *Automatica*, vol. 107, pp. 9–20, 2019.
- [5] S. Saritaş, S. Yüksel, and S. Gezici, "Dynamic signaling games with quadratic criteria under Nash and Stackelberg equilibria," *Automatica*, vol. 115, p. 108883, 2020.
- [6] S. Saritaş, S. Gezici, and S. Yüksel, "Hypothesis testing under subjective priors and costs as a signaling game," *IEEE Transactions on Signal Processing*, vol. 67, no. 19, pp. 5169–5183, 2019.
- [7] M. L. Treust and T. Tomala, "Strategic communication with side information at the decoder," *arXiv preprint arXiv:1911.04950*, Sept. 2020.
- [8] M. Le Treust and T. Tomala, "Persuasion with limited communication capacity," *Journal of Economic Theory*, vol. 184, p. 104940, 2019.
- [9] A. S. Vora and A. A. Kulkarni, "Information extraction from a strategic sender over a noisy channel," in *IEEE Conference on Decision and Control (CDC)*, pp. 354–359, 2020.
- [10] A. S. Vora and A. A. Kulkarni, "Achievable rates for strategic communication," in *IEEE International Symposium on Information Theory (ISIT)*, pp. 1379–1384, 2020.
- [11] E. Kazıklı, S. Saritaş, S. Gezici, T. Linder, and S. Yüksel, "Signaling games for arbitrary distributions: Number of bins and properties of equilibria," *arXiv preprint arXiv:2012.08265*, Dec. 2020.
- [12] E. Kazıklı, S. Saritaş, S. Gezici, and S. Yüksel, "Optimal signaling with mismatch in priors of an encoder and decoder," *arXiv preprint arXiv:2101.00799*, Jan. 2021.
- [13] T. Başar, "An equilibrium theory for multiperson decision making with multiple probabilistic models," *IEEE Transactions on Automatic Control*, vol. 30, pp. 118–132, Feb. 1985.
- [14] V. P. Crawford and J. Sobel, "Strategic information transmission," *Econometrica*, vol. 50, no. 6, pp. 1431–1451, 1982.
- [15] T. Başar and G. J. Olsder, *Dynamic Noncooperative Game Theory*. Philadelphia, PA: SIAM Classics in Applied Mathematics, 1999.
- [16] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. New Jersey: John Wiley & Sons, 2006.
- [17] J. Sobel, *Signaling Games*, pp. 251–268. New York, NY: Springer US, 2020.
- [18] E. Kamenica and M. Gentzkow, "Bayesian persuasion," *American Economic Review*, vol. 101, pp. 2590–2615, Oct. 2011.
- [19] B. Larrousse, O. Beaude, and S. Lasaulce, "Crawford-Sobel meet Lloyd-Max on the grid," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 6127–6131, 2014.
- [20] V. Kavitha, E. Altman, R. El-Azouzi, and R. Sundaresan, "Opportunistic scheduling in cellular systems in the presence of noncooperative mobiles," *IEEE Transactions on Information Theory*, vol. 58, no. 3, pp. 1757–1773, 2012.
- [21] A. S. Vora and A. A. Kulkarni, "Optimal questionnaires for screening of strategic agents," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 8173–8177, 2021.
- [22] G. Levy and R. Razin, "On the limits of communication in multidimensional cheap talk: A comment," *Econometrica*, vol. 75, no. 3, pp. 885–893, 2007.
- [23] M. Battaglini, "Multiple referrals and multidimensional cheap talk," *Econometrica*, vol. 70, no. 4, pp. 1379–1401, 2002.
- [24] S. Miura, "Multidimensional cheap talk with sequential messages," *Games and Economic Behavior*, vol. 87, pp. 419–441, 2014.
- [25] E. Kazıklı, S. Gezici, and S. Yüksel, "Signaling games in higher dimensions: Geometric properties of equilibrium solutions," *arXiv preprint arXiv:2108.05240*, Aug. 2021.
- [26] A. M. Kagan, Y. V. Linnik, and C. R. Rao, *Characterization Problems in Mathematical Statistics*. New York: John Wiley & Sons, 1973.
- [27] J. C. Harsanyi and R. Selten, *A General Theory of Equilibrium Selection in Games*. Cambridge, Massachusetts: MIT Press, 1988.