Quadratic Privacy-Signaling Games and Payoff Dominant Equilibria

Ertan Kazıklı
Dept. of Electrical and Electronics Eng.
Bilkent University
06800, Ankara, Turkey
kazikli@ee.bilkent.edu.tr

Sinan Gezici
Dept. of Electrical and Electronics Eng.
Bilkent University
06800, Ankara, Turkey
gezici@ee.bilkent.edu.tr

Serdar Yüksel
Dept. of Mathematics and Statistics
Queen's University
Kingston, Ontario, Canada, K7L 3N6
yuksel@mast.queensu.ca

***Abstract*—We consider a privacy-signaling game problem in which a transmitter with privacy concerns and a receiver, which does not pay attention to these privacy concerns, communicate. In this communication scenario, the transmitter observes a pair of correlated random variables which are modeled as jointly Gaussian. The transmitter constructs its message based on these random variables with the aim to hide one of them and convey the other one. In contrast, the objective of the receiver is to accurately estimate both of the random variables so as to gather as much information as possible. These conflicting objectives are analyzed in a game theoretic framework where depending on the commitment conditions (of the sender), we consider Nash or Stackelberg equilibria. We show that a payoff dominant (i.e., most desirable for both players) Nash equilibrium is attained by affine policies and we explicitly characterize these policies. In addition, the strategies at the characterized Nash equilibrium is shown to form also a Stackelberg equilibrium. Furthermore, we show that there always exists an informative Stackelberg equilibrium for the multidimensional parameter setup. We also revisit the information bottleneck problem within our Stackelberg framework under the mean squared error distortion criterion where the information bottleneck setup has a further restriction that only one of the parameters is observed at the sender. We fully characterize the Stackelberg equilibria under certain conditions and when these conditions are not met we establish the existence of informative equilibria.**

I. Introduction and System Model

We consider the following communication scenario between a sender and a receiver motivated by various applications such as smart grid and crowd sensing [1]–[8]. There is a pair of messages at the sender and the perspective of the sender is such that one message needs to be protected and the other message needs to be conveyed. As opposed to the sender, the receiver desires to accurately estimate both messages with the aim of acquiring as much information as possible. The problem considered in this manuscript belongs to a general class of problems referred to as strategic information transmission (SIT) problems where communication scenarios between agents with misaligned objectives are investigated.

Consider an information transmission scenario in which a sender encodes a pair of correlated random variables x and y into z using an encoding function denoted by $z = \gamma^e(x, y)$ and a receiver wants to decode both of the random variables. In this communication scenario, the sender wishes to convey information contained in y whereas it views x as a private parameter that needs to be hidden from the receiver. The aim of the receiver is to accurately estimate both parameters given its observation z. Let the decoders for estimating x and y at the receiver be denoted by $\gamma^{d_X}(z)$ and $\gamma^{d_Y}(z)$, respectively. Under this setting, we assume that there is a power constraint $\mathbb{E}[z^2] \leq P$ at the transmitter where P represents the maximum average power.

We model the parameters x and y as jointly Gaussian distributed random variables. Let (x, y) be a zero mean Gaussian distributed random vector with a positive definite covariance matrix $\Sigma \triangleq \begin{bmatrix} \sigma_X^2 & \rho \\ \rho & \sigma_Y^2 \end{bmatrix}$ and a nonzero correlation, i.e., $\rho \neq 0$. It is assumed that the joint distribution of x and y is common knowledge, i.e., both players know σ_X^2, σ_Y^2 and ρ. Since x and y are correlated, transmitting y directly discloses information related to the private parameter x. In other words, the objectives of hiding x and conveying y are conflicting. These conflicting objectives at the transmitter are modeled via the following objective function:

$$J^e(\gamma^e, \gamma^{d_X}, \gamma^{d_Y}) = \mathbb{E}[(\gamma^{d_Y}(z) - y)^2] - \delta\, \mathbb{E}[(\gamma^{d_X}(z) - x)^2] \tag{1}$$

which is to be minimized, where δ is a positive design parameter that determines the level of desired privacy in terms of hiding x. On the other hand, the receiver aims to extract both of the parameters. Thus, the receiver wishes to minimize the following objective function:

$$J^d(\gamma^e, \gamma^{d_X}, \gamma^{d_Y}) = \mathbb{E}[(\gamma^{d_X}(z) - x)^2] + \mathbb{E}[(\gamma^{d_Y}(z) - y)^2]. \tag{2}$$

In this work, we investigate the Nash and the Stackelberg equilibria for the described strategic information transmission scenario in which the objectives of the sender and the receiver are as defined above.

The game dynamics for the Nash equilibrium are as follows. Each player chooses its strategy simultaneously. These chosen strategies are referred to as a Nash equilibrium if no player gains by unilaterally deviating from its strategy. In other words, neither the sender nor the receiver would have any incentive to unilaterally change their strategies when they operate at a Nash equilibrium. Suppose that the set of possible strategies for the encoder is denoted by Γ^e, i.e., $\gamma^e \in \Gamma^e$,

and those for the decoders of each parameter are denoted by Γ^{d_X} and Γ^{d_Y}, i.e., $\gamma^{d_X} \in \Gamma^{d_X}$ and $\gamma^{d_Y} \in \Gamma^{d_Y}$. A set of policies $\gamma^{e,*}$, $\gamma^{d_X,*}$ and $\gamma^{d_Y,*}$ forms a Nash equilibrium if $J^e(\gamma^{e,*}, \gamma^{d_X,*}, \gamma^{d_Y,*}) \leq J^e(\gamma^e, \gamma^{d_X,*}, \gamma^{d_Y,*})$ for all $\gamma^e \in \Gamma^e$ and $J^d(\gamma^{e,*}, \gamma^{d_X,*}, \gamma^{d_Y,*}) \leq J^d(\gamma^{e,*}, \gamma^{d_X}, \gamma^{d_Y})$ for all $\gamma^{d_X} \in \Gamma^{d_X}$ and $\gamma^{d_Y} \in \Gamma^{d_Y}$.

On the other hand, the Stackelberg equilibrium involves a sequential game play in the sense that first the sender and then the receiver acts. The sender chooses and announces its strategy and then the receiver acts upon learning the strategy of the sender. Once the sender announces its strategy, it cannot change it, i.e., the sender commits to employ this announced strategy. The receiver employs an optimal response to the announced strategy of the sender, which is known by the receiver before taking an action. A set of policies $\gamma^{e,*}$, $\gamma^{d_X,*}$ and $\gamma^{d_Y,*}$ forms a Stackelberg equilibrium if

$$J^e(\gamma^{e,*}, \gamma^{d_X,*}(\gamma^{e,*}), \gamma^{d_Y,*}(\gamma^{e,*})) \leq J^e(\gamma^e, \gamma^{d_X,*}(\gamma^e), \gamma^{d_Y,*}(\gamma^e))$$

for all $\gamma^e \in \Gamma^e$, where $\gamma^{d_X,*}(\gamma^e)$ and $\gamma^{d_Y,*}(\gamma^e)$ are such that

$$J^d(\gamma^e, \gamma^{d_X,*}(\gamma^e), \gamma^{d_Y,*}(\gamma^e)) \leq J^d(\gamma^e, \gamma^{d_X}(\gamma^e), \gamma^{d_Y}(\gamma^e))$$

for all $\gamma^{d_X} \in \Gamma^{d_X}$ and $\gamma^{d_Y} \in \Gamma^{d_Y}$.

II. Literature Review

For signaling games under the Nash equilibrium concept, Crawford and Sobel in their foundational paper [9] investigate a communication scenario between a better informed sender and a receiver where sender's cost contains a bias term leading to misaligned objectives. They obtain the interesting result that under some technical conditions the sender needs to quantize the information it sends at a Nash equilibrium. To put it differently, the misalignment in the objectives results in information hiding through quantization of the sent message. In contrast to Crawford and Sobel, the Bayesian persuasion problem considers the Stackelberg equilibrium concept rather than the Nash equilibrium concept [10].

In the literature, several studies consider the SIT problem in which the sender takes privacy of certain information into account by employing a suitable privacy measure, under either the Nash or Stackelberg criteria [11]–[14]. In these studies, a common theme is to model private and nonprivate parameters as jointly Gaussian random variables. In [11], a communication scenario between a sender and a receiver is investigated using the Stackelberg equilibrium concept in which an additional side information is assumed to be available at the receiver. The estimation errors are measured using quadratic costs and a family of Stackelberg equilibria is characterized under an a priori affine policy assumption. In contrast, here, we do not restrict the policies to be affine a priori and we consider a setting with no side information. Also, we investigate Nash equilibria as well and show that a payoff dominant equilibrium is attained by affine policies. We use this result to show that the Stackelberg equilibria are also affine even when the sender is not restricted to the affine class. The work in [12] also investigates a Stackelberg game where the utility measure for the nonprivate parameter is quadratic and the privacy measure is entropy based. In [13], a Nash game is studied where the privacy measure is based on mutual information and the utility measure for the nonprivate parameter is quadratic.

The tradeoff between utility and privacy appears also in various other contexts [15]–[23]. One important related line of work is the information bottleneck technique where the aim is to compress an observed random variable while trying to preserve information related to another correlated random variable which is not observed [24], [25]. The compression objective in the information bottleneck problem can also be viewed as a privacy objective as in our framework in the sense that the corresponding information is desired to be removed from the revealed message. In the information bottleneck problem, the costs involve mutual information and only one of the parameters is received at the sender whereas in our framework the costs include mean squared error terms and both of the parameters are observed at the sender. In order to provide an estimation theoretic perspective on the information bottleneck problem, we formulate a similar problem where we use mean squared error terms for the costs as in our original setting and we show that there are operational and consequential differences when the encoder is allowed to use both of the hidden variables.

For a more detailed literature review, please see [26].

III. Nash Equilibria

In the following theorem, we show that payoff dominant Nash equilibria are affine and we explicitly characterize these equilibria. A payoff dominant Nash equilibrium is the most desirable equilibrium for both of the players (among all coding/decoding policies, including those that are non-linear) in a sense that is made explicit in the following definition [27].

Definition III.1 *A Nash equilibrium that is not Pareto dominated*[1] *by any other Nash equilibrium of the game is said to be a payoff dominant Nash equilibrium.*

Theorem III.1 (i) *The encoding policies $\gamma^e(x,y) = C$ for all $|C| \leq \sqrt{P}$ and the decoding policy $\gamma^{d_X}(z) = 0$ and $\gamma^{d_Y}(z) = 0$ form a noninformative Nash equilibrium.*[2]

(ii) *There exist informative affine Nash equilibria in which the decoding policies, $\gamma^{d_X}(z) = Kz + L$ and $\gamma^{d_Y}(z) = Mz + N$, satisfy*

$$\frac{M}{K} = \frac{\delta\sigma_x^2 + \sigma_y^2}{2\rho} + \frac{\sqrt{(\delta\sigma_x^2 + \sigma_y^2)^2 - 4\delta\rho^2}}{2\rho} \qquad (3)$$

[1] A set of policies $\gamma^e(\cdot,\cdot)$, $\gamma^{d_X}(\cdot)$ and $\gamma^{d_Y}(\cdot)$ Pareto dominates another set of policies $\tilde{\gamma}^e(\cdot,\cdot)$, $\tilde{\gamma}^{d_X}(\cdot)$ and $\tilde{\gamma}^{d_Y}(\cdot)$ if $J^e(\gamma^e, \gamma^{d_X}, \gamma^{d_Y}) \leq J^e(\tilde{\gamma}^e, \tilde{\gamma}^{d_X}, \tilde{\gamma}^{d_Y})$, $J^d(\gamma^e, \gamma^{d_X}, \gamma^{d_Y}) \leq J^d(\tilde{\gamma}^e, \tilde{\gamma}^{d_X}, \tilde{\gamma}^{d_Y})$ and at least one of these inequalities is strict.

[2] We refer to an equilibrium as noninformative if the sender does not convey information related to both of the parameters at this equilibrium and this is equivalent to what is known as a *babbling equilibrium* in the signaling games literature. In the converse case, the equilibrium is referred to as informative.

with $\tilde{P}(K,L,M,N) \leq (M^2-\delta K^2)^2 P$, and L and N are such that either $N/L = M/K$ or $L = N = 0$ holds. Also, the corresponding encoding policy, $\gamma^e(x,y) = Ax + By + C$, is given by $A = -\delta K/(M^2-\delta K^2)$, $B = M/(M^2-\delta K^2)$ and $C = -L/K$. In addition, these equilibria are payoff dominant Nash equilibria, and they are unique among the affine class of policies.

Proof Sketch: In the rest of this section, we prove Theorem III.1, where we focus on (ii). We first characterize affine Nash equilibria and then we prove their payoff dominance property among all encoding and decoding policies, including those that are non-linear. Towards that goal, we first need the best responses of each player specializing to affine policies.

Lemma III.1 *When the encoder is affine and it is expressed as $z = Ax + By + C$, the decoders are also affine and they are given by*

$$\gamma^{d_X}(z) = \left(\frac{A\sigma_X^2 + B\rho}{A^2\sigma_X^2 + B^2\sigma_Y^2 + 2AB\rho}\right)(z - C), \quad (4)$$

$$\gamma^{d_Y}(z) = \left(\frac{A\rho + B\sigma_Y^2}{A^2\sigma_X^2 + B^2\sigma_Y^2 + 2AB\rho}\right)(z - C). \quad (5)$$

Proof Sketch: The proof uses [28, p. 155]. ■

Lemma III.2 *Define $\tilde{P}(K,L,M,N) \triangleq \delta^2\sigma_X^2K^2 + \sigma_Y^2M^2 - 2\delta\rho KM + (MN - \delta KL)^2$. When the decoders are affine in the form of $\gamma^{d_X}(z) = Kz + L$ and $\gamma^{d_Y}(z) = Mz + N$, the optimal encoder is specified by*

$$\gamma^e(x,y) = \frac{M(y-N) - \delta K(x-L)}{M^2 - \delta K^2} \quad (6)$$

if $\tilde{P}(K,L,M,N) \leq (M^2-\delta K^2)^2 P$ and $M^2 > \delta K^2$; and otherwise by

$$\gamma^e(x,y) = \frac{\sqrt{P}M\,(y-N) - \sqrt{P}\delta K\,(x-L)}{\sqrt{\tilde{P}(K,L,M,N)}}. \quad (7)$$

Proof Sketch: The proof essentially uses Cauchy-Schwarz inequality and is omitted due to space limitations. ■

Suppose that the system parameters and the terms that specify the decoding policies are such that the optimal encoding policy is as in (6). Namely, it is assumed that $M^2 > \delta K^2$ and $\tilde{P}(K,L,M,N) \leq P(M^2-\delta K^2)^2$ hold. In the following, a Nash equilibrium analysis is carried out under these two assumptions. By combining (4)-(6), manipulating the resulting fixed point equations and defining $q \triangleq M/K$, we need $q^2\rho - q(\delta\sigma_x^2 + \sigma_y^2) + \delta\rho = 0$. For this to be solvable, it is required that $(\delta\sigma_x^2 + \sigma_y^2)^2 - 4\delta\rho^2 \geq 0$. By using positive definiteness property of the covariance matrix Σ, it can be shown that this condition is met with strict inequality. Then, it can be established that the root given in the theorem statement satisfies $M^2 > \delta K^2$ whereas the other root does not. The encoding policy at this equilibrium is as in (6) which gives the encoding policy stated in the theorem.

Next, we propose an equivalent formulation to show the payoff dominance property of the derived affine Nash equilibria as well as its uniqueness. Towards that goal, we linearly transform (x,y) into (u,v) such that u corresponds to the revealed parameter at previously characterized equilibria and v is independent of u. Note that this transformation is invertible. In particular, consider the following transformation:

$$u = \left(-\frac{\delta\sigma_X^2 + \sigma_Y^2}{2\delta\rho} - \frac{\sqrt{(\delta\sigma_X^2 + \sigma_Y^2)^2 - 4\delta\rho^2}}{2\delta\rho}\right)y + x, \quad (8)$$

$$v = \left(-\frac{\delta\sigma_X^2 + \sigma_Y^2}{2\delta\rho} + \frac{\sqrt{(\delta\sigma_X^2 + \sigma_Y^2)^2 - 4\delta\rho^2}}{2\delta\rho}\right)y + x. \quad (9)$$

In the equivalent problem, there is a linear mapping from (x,y) to (u,v) which is fixed as above and then an encoding function $z = \tilde{\gamma}^e(u,v)$ which can arbitrarily be chosen by the sender. At the receiver side, the observation is mapped into $\tilde{u}$ and $\tilde{v}$ via $\gamma^{d_U}(z)$ and $\gamma^{d_V}(z)$, respectively, which can arbitrarily be selected by the receiver. Then, these auxiliary variables are mapped into estimates of x and y as follows:

$$\gamma^{d_X}(z) = \frac{\gamma^{d_U}(z) + \gamma^{d_V}(z)}{2} + \frac{(\gamma^{d_V}(z) - \gamma^{d_U}(z))(\delta\sigma_X^2 + \sigma_Y^2)}{2\sqrt{(\delta\sigma_X^2 + \sigma_Y^2)^2 - 4\delta\rho^2}}, \quad (10)$$

$$\gamma^{d_Y}(z) = \frac{(\gamma^{d_V}(z) - \gamma^{d_U}(z))\delta\rho}{\sqrt{(\delta\sigma_X^2 + \sigma_Y^2)^2 - 4\delta\rho^2}}. \quad (11)$$

If we express the objective function of the sender in terms of the parameters in the introduced coordinate system, we get

$$J^e(\tilde{\gamma}^e, \gamma^{d_U}, \gamma^{d_V}) = \frac{\delta}{2}\left(\frac{(\delta\sigma_X^2 + \sigma_Y^2)}{\sqrt{(\delta\sigma_X^2 + \sigma_Y^2)^2 - 4\delta\rho^2}} - 1\right)\mathbb{E}[(u - \gamma^{d_U}(z))^2] + \frac{\delta}{2}\left(-\frac{(\delta\sigma_X^2 + \sigma_Y^2)}{\sqrt{(\delta\sigma_X^2 + \sigma_Y^2)^2 - 4\delta\rho^2}} - 1\right)\mathbb{E}[(v - \gamma^{d_V}(z))^2], \quad (12)$$

where the first coefficient is positive and the second coefficient is negative. Similarly, the objective function of the receiver becomes $J^d(\tilde{\gamma}^e, \gamma^{d_U}, \gamma^{d_V}) = w_1\mathbb{E}[(u-\gamma^{d_U}(z))^2] + w_2\mathbb{E}[(v-\gamma^{d_V}(z))^2] + w_3\mathbb{E}[(u-\gamma^{d_U}(z))(v-\gamma^{d_V}(z))]$ where w_1, w_2 and w_3 are known coefficients. By using the completion of squares technique, it can be shown that for a fixed encoding function, the optimal $\gamma^{d_U}(z)$ and $\gamma^{d_V}(z)$ are given by $\mathbb{E}[u|z]$ and $\mathbb{E}[v|z]$, respectively.

Lemma III.3 *At a Nash equilibrium, the sender does not transmit any information related to the linear combination (of the private and nonprivate parameters) v specified by* (9)*.*

Proof Sketch: Suppose that the sender employs an encoding policy $\tilde{\gamma}^e(u,v)$ so that the mean squared error for estimating v with the corresponding optimal estimator is less than σ_V^2. In response to the optimal estimators employed by the receiver, the sender can switch to the following policy to

improve its objective value. Instead of sending $z = \tilde{\gamma}^e(u, v)$, the sender can transmit $z = \tilde{\gamma}^e(u, n)$ while keeping the encoding function $\tilde{\gamma}^e(\cdot, \cdot)$ the same where n follows the same distribution as v and is independent of u and v. In that case, the performance for estimating u remains the same since receiving $\tilde{\gamma}^e(u, v)$ or $\tilde{\gamma}^e(u, n)$ are equivalent. However, it can be shown that the performance for estimating v degrades. Since the sender wishes to hide v (see (12)), the sender gains by employing $z = \tilde{\gamma}^e(u, n)$ instead of $z = \tilde{\gamma}^e(u, v)$. ■

From Lemma III.3, it follows that the performance of each player at a Nash equilibrium is determined by the mean squared error for estimating u and this result is valid for any coding/decoding policies including non-linear ones. Since transmitting an affine function of u yields the minimum attainable performance for both players, this is a payoff dominant Nash equilibrium, i.e., the most desirable equilibrium for both players among all coding/decoding policies including non-linear ones. In addition, these equilibria are unique among the affine class since the sender is restricted to send u at a Nash equilibrium. ■

IV. Stackelberg Equilibria

Lemma IV.1 *At a Stackelberg equilibrium, the sender does not transmit any information related to the linear combination (of the private and nonprivate parameters) v specified by* (9)*.*

Proof Sketch: The result can be proven via a similar analysis to that employed in Lemma III.3 with the exception that the sender announces its strategy and commits to this announced strategy under the Stackelberg equilibrium concept. ■

Theorem IV.1 *The Stackelberg equilibria coincide with the informative Nash equilibria characterized in Theorem III.1. In addition, these equilibria are unique among the affine class of policies.*

It is important to emphasize that the set of possible encoding strategies, i.e., Γ^e, is not restricted to the affine class.

V. The Multidimensional Case

A. Nash Equilibria

Here, the private and nonprivate parameters as well as the message to be transmitted are multidimensional. We impose no power constraint at the sender to simplify the problem. Similar to the scalar setting, the objective functions of the sender and receiver are defined as

$$J^e(\gamma^e, \gamma^{d_{\boldsymbol{X}}}, \gamma^{d_{\boldsymbol{Y}}}) = -\delta\mathbb{E}[\|\boldsymbol{x} - \hat{\boldsymbol{x}}\|^2] + \mathbb{E}[\|\boldsymbol{y} - \hat{\boldsymbol{y}}\|^2], \quad (13)$$

$$J^d(\gamma^e, \gamma^{d_{\boldsymbol{X}}}, \gamma^{d_{\boldsymbol{Y}}}) = \mathbb{E}[\|\boldsymbol{x} - \hat{\boldsymbol{x}}\|^2] + \mathbb{E}[\|\boldsymbol{y} - \hat{\boldsymbol{y}}\|^2]. \quad (14)$$

Theorem V.1 *Consider linear decoding policies $\gamma^{d_{\boldsymbol{X}}}(\boldsymbol{z}) = K\boldsymbol{z}$ and $\gamma^{d_{\boldsymbol{Y}}}(\boldsymbol{z}) = M\boldsymbol{z}$. If a set of linear decoding policies with $M^TM - \delta K^TK$ being positive definite and, M and K satisfying*

$$K = \Big(\Sigma_{\boldsymbol{XY}}M - \delta\Sigma_{\boldsymbol{XX}}K\Big)\Big(\delta^2K^T\Sigma_{\boldsymbol{XX}}K + M^T\Sigma_{\boldsymbol{YY}}M - \delta K^T\Sigma_{\boldsymbol{XY}}M - \delta M^T\Sigma_{\boldsymbol{YX}}K\Big)^{-1}\Big(M^TM - \delta K^TK\Big),$$

$$M = \Big(\Sigma_{\boldsymbol{YY}}M - \delta\Sigma_{\boldsymbol{YX}}K\Big)\Big(\delta^2K^T\Sigma_{\boldsymbol{XX}}K + M^T\Sigma_{\boldsymbol{YY}}M - \delta K^T\Sigma_{\boldsymbol{XY}}M - \delta M^T\Sigma_{\boldsymbol{YX}}K\Big)^{-1}(M^TM - \delta K^TK),$$

exists, then these linear policies yield a Nash equilibrium assuming that the sender employs its best response.

We note that finding an informative equilibrium or showing its existence is challenging since the standard approaches of applying fixed point theorems such as Brouwer's [29] is not directly applicable as there is always a noninformative equilibrium. In the following theorem, we explicitly characterize Nash equilibria considering special covariance matrix structures.

Theorem V.2 *Under either of the following cases, there exist linear Nash equilibria, which are payoff dominant equilibria.*

(i) *For independent pairs of observations $(x_1, y_1), \ldots, (x_n, y_n)$, transmitting their pairwise linear combinations $u_1, \ldots, u_n$, where u_i is as in* (8) *with parameters x_i and y_i, yields a payoff dominant Nash equilibrium.*

(ii) *Consider pairs of observations $(x_1, y_1), \ldots, (x_n, y_n)$, where $x_i = \alpha_i x_1$ and $y_i = \alpha_i y_1$ for $i = 2, \ldots, n$ and nonzero scalars α_i. Then, transmitting the linear combination u specified by* (8) *with parameters x_1 and y_1 yields a payoff dominant Nash equilibrium.*

B. Stackelberg Equilibria

In this section, our main result is to show that there always exists an informative Stackelberg equilibrium. Towards that goal, we provide a lower bound for the performance of the encoder under which the encoder conveys information related to both parameters. This implies that the Stackelberg equilibrium should be informative since the Stackelberg equilibrium cannot yield a performance which is worse than this informative lower bound. In the following, we provide an analysis building on the approach initiated in [11] where the authors consider a slightly different setting in which side information is available at the receiver.

Theorem V.3 *There always exists an informative Stackelberg equilibrium for the multidimensional parameter setting.*

Proof Sketch: Suppose that the receiver employs the following linear policies $\gamma^{d_{\boldsymbol{X}}}(\boldsymbol{z}) = \Sigma_{\boldsymbol{XZ}}\Sigma_{\boldsymbol{ZZ}}^{-1}\boldsymbol{z}$ and $\gamma^{d_{\boldsymbol{Y}}}(\boldsymbol{z}) = \Sigma_{\boldsymbol{YZ}}\Sigma_{\boldsymbol{ZZ}}^{-1}\boldsymbol{z}$, which become the best response of the receiver when the sender employs linear policies. Then, the mean squared errors for estimating each parameter can be written as $\mathbb{E}[\|\boldsymbol{x} - \Sigma_{\boldsymbol{XZ}}\Sigma_{\boldsymbol{ZZ}}^{-1}\boldsymbol{z}\|] = \mathrm{Tr}(\Sigma_{\boldsymbol{XX}} - \Sigma_{\boldsymbol{ZX}}\Sigma_{\boldsymbol{XZ}}\Sigma_{\boldsymbol{ZZ}}^{-1})$ and $\mathbb{E}[\|\boldsymbol{y} - \Sigma_{\boldsymbol{YZ}}\Sigma_{\boldsymbol{ZZ}}^{-1}\boldsymbol{z}\|] = \mathrm{Tr}(\Sigma_{\boldsymbol{YY}} - \Sigma_{\boldsymbol{ZY}}\Sigma_{\boldsymbol{YZ}}\Sigma_{\boldsymbol{ZZ}}^{-1})$. Without loss of generality, we assume that $\Sigma_{\boldsymbol{ZZ}} = I$ since it is possible to scale the message without changing the objective. Since we are looking for a lower bound, let us assume that z is scalar.

Thus, after defining $\boldsymbol{u} \triangleq \begin{bmatrix}\Sigma_{\boldsymbol{X}Z}\\ \Sigma_{\boldsymbol{Y}Z}\end{bmatrix}$ and $Q \triangleq \begin{bmatrix}\Sigma_{\boldsymbol{XX}} & \Sigma_{\boldsymbol{XY}}\\ \Sigma_{\boldsymbol{YX}} & \Sigma_{\boldsymbol{YY}}\end{bmatrix}$, the optimization problem at the sender can be written as

$$\min_{\boldsymbol{u}} \boldsymbol{u}^T \mathrm{diag}(\delta \mathrm{I}, -\mathrm{I})\boldsymbol{u}, \quad \text{subject to}\, \boldsymbol{u}^T Q^{-1}\boldsymbol{u} \leq 1, \tag{15}$$

where the constraint is due to positive semi-definiteness condition expressed using Schur's complement. If we write $Q = P^2$ and apply a transformation of variables, we obtain an equivalent optimization problem

$$\min_{\tilde{\boldsymbol{u}}} \tilde{\boldsymbol{u}}^T P \mathrm{diag}(\delta I, -I) P \tilde{\boldsymbol{u}}, \quad \text{subject to}\, \tilde{\boldsymbol{u}}^T \tilde{\boldsymbol{u}} \leq 1. \tag{16}$$

We can show that the constraint in (16) should be satisfied with equality. Then, by Courant-Fischer-Weyl minimax principle [30, p.58], the optimal solution is given by the normalized eigenvector corresponding to the smallest eigenvalue of $P\mathrm{diag}(\delta I, -I)P$. Since employing a linear encoding policy that achieves optimal solution of (16) gives an informative lower bound, it follows that the Stackelberg equilibrium should be informative. ■

We note that this characterization gives a lower bound for the Stackelberg equilibrium since there is a linear policy restriction as well as a scalar message restriction.

VI. Information Bottleneck interpreted with the paper's formulation

We now revisit the information bottleneck problem [24], [25], which is a popular framework in the current literature, as an instance of our formulation under the Stackelberg equilibrium concept in the following sense: in contrast to privacy game setup, only the parameter $\boldsymbol{x}$ is observed at the sender in the information bottleneck setup. In the information bottleneck problem, the costs involve mutual information leading to an information theoretic framework. The use of mutual information effectively means that the receiver uses all the information available, i.e., it employs its best response. Thus, the information bottleneck problem can in fact be viewed as a Stackelberg game between a sender and a receiver. In this manuscript, we provide an estimation theoretic perspective on the information bottleneck problem by using mean squared error as our metric.

As in the information bottleneck framework, the sender observes only the parameter $\boldsymbol{x}$, rather than observing both parameters. The objective functions of the sender and receiver are as defined in (13) and (14), respectively. Since the receiver is concerned with estimating both parameters, it employs the minimum mean squared error estimators of each parameter. Since the equilibrium concept is the Stackelberg equilibrium, the objective function of the sender can be written as $J^e(\gamma^e) = -\delta\,\mathbb{E}[\|\boldsymbol{x} - \mathbb{E}[\boldsymbol{x}|\boldsymbol{z}]\|^2] + \mathbb{E}[\|\boldsymbol{y} - \mathbb{E}[\boldsymbol{y}|\boldsymbol{z}]\|^2]$. Now, observe that

$$\begin{aligned}\mathbb{E}[\|\boldsymbol{y} - \mathbb{E}[\boldsymbol{y}|\boldsymbol{z}]\|^2] &= \mathbb{E}[\|\boldsymbol{y} - \mathbb{E}[\boldsymbol{y}|\boldsymbol{x}] + \mathbb{E}[\boldsymbol{y}|\boldsymbol{x}] - \mathbb{E}[\boldsymbol{y}|\boldsymbol{z}]\|^2]\\ &= \mathbb{E}[\|\boldsymbol{y} - \mathbb{E}[\boldsymbol{y}|\boldsymbol{x}]\|^2] + \mathbb{E}[\|\mathbb{E}[\boldsymbol{y}|\boldsymbol{x}] - \mathbb{E}[\boldsymbol{y}|\boldsymbol{z}]\|^2]\\ &= \mathbb{E}[\|\boldsymbol{y} - \mathbb{E}[\boldsymbol{y}|\boldsymbol{x}]\|^2] + \mathbb{E}[\|\mathbb{E}[\boldsymbol{y}|\boldsymbol{x}] - \mathbb{E}[\mathbb{E}[\boldsymbol{y}|\boldsymbol{x}]|\boldsymbol{z}]\|^2]\end{aligned} \tag{17}$$

where the second equality follows from the fact that $\boldsymbol{y} - \mathbb{E}[\boldsymbol{y}|\boldsymbol{x}]$ is orthogonal to $\boldsymbol{x}$ and $\boldsymbol{z}$ (orthogonality with respect to $\boldsymbol{z}$ is due to the fact that $\boldsymbol{y} - \boldsymbol{x} - \boldsymbol{z}$ is a Markov chain in that order) and the third equality is due to the law of iterated expectations. By noticing that $\mathbb{E}[\boldsymbol{y}|\boldsymbol{x}] = \Sigma_{\boldsymbol{YX}}\Sigma_{\boldsymbol{X}}^{-1}\boldsymbol{x}$, we obtain an optimization problem of the form

$$\min_{\boldsymbol{z}=\gamma^e(\boldsymbol{x})} \mathbb{E}\Big[(\boldsymbol{x} - \mathbb{E}[\boldsymbol{x}|\boldsymbol{z}])^T M (\boldsymbol{x} - \mathbb{E}[\boldsymbol{x}|\boldsymbol{z}])\Big], \tag{18}$$

where $M \triangleq (\Sigma_{\boldsymbol{X}}^{-1}\Sigma_{\boldsymbol{XY}}\Sigma_{\boldsymbol{YX}}\Sigma_{\boldsymbol{X}}^{-1} - \delta I)$. After observing that M is symmetric, we can write $M = Q\Lambda Q^T$ where Λ is a diagonal matrix and Q is an orthonormal matrix. If we define $\tilde{\boldsymbol{x}} \triangleq Q^T\boldsymbol{x}$, we obtain an equivalent problem:

$$\min_{\boldsymbol{z}=\gamma^e(\boldsymbol{x})} \mathbb{E}\Big[\mathrm{tr}\Big(\Lambda(\tilde{\boldsymbol{x}} - \mathbb{E}[\tilde{\boldsymbol{x}}|\boldsymbol{z}])(\tilde{\boldsymbol{x}} - \mathbb{E}[\tilde{\boldsymbol{x}}|\boldsymbol{z}])^T\Big)\Big]. \tag{19}$$

Under certain symmetry conditions, it is possible to solve the optimization problem in (19). Let λ_i denote the ith diagonal element of Λ and $\tilde{x}_i$ denote the ith element of $\tilde{\boldsymbol{x}}$.

Theorem VI.1 (i) *If the eigenvectors of M and $\Sigma_{\boldsymbol{X}}$ are completely aligned, then the sender transmits all $\tilde{x}_i$ for which $\lambda_i \geq 0$ at the Stackelberg equilibrium.*

(ii) *If M is positive semi-definite, then there exists a fully informative Stackelberg equilibrium in which the sender transmits $\boldsymbol{x}$.*

(iii) *If M is negative definite, then the equlibrium is always noninformative.*

(iv) *Let M be negative semi-definite with at least one eigenvalue equal to zero. Then, there exists an informative Stackelberg equilibrium if a component $\tilde{x}_i$ for which $\lambda_i = 0$ is uncorrelated with any component $\tilde{x}_j$ for which $\lambda_j < 0$. Otherwise, the equilibrium is always noninformative.*

Theorem VI.1 implies that for scalar sources the equilibrium is informative if $\delta < (\rho^2/\sigma_X^4)$ and noninformative otherwise.

For the case when M is neither positive semi-definite nor negative semi-definite, the reduced problem in (19) can be viewed as a multidimensional privacy-signaling game under the Stackelberg equilibrium concept by appropriately partitioning $\tilde{\boldsymbol{x}}$ into private and nonprivate parameters based on the sign of diagonal terms of Λ. The following result uses this observation and Theorem V.3.

Theorem VI.2 *If M has at least one positive and at least one negative eigenvalue, then there exists an informative Stackelberg equilibrium.*

VII. Conclusion

A communication setting between a sender with privacy concerns and a receiver has been investigated in a game theoretic framework. The private and nonprivate parameters have been modeled as jointly Gaussian random variables. For this privacy-signaling game formulation, the Nash and the Stackelberg equilibria have been investigated. Moreover, a Gaussian information bottleneck problem under the mean squared error distortion criterion and under the Stackelberg equilibrium concept has been formulated and equilibrium solutions have been investigated.

REFERENCES

[1] S. Finster and I. Baumgart, "Privacy-aware smart metering: A survey," *IEEE Communications Surveys Tutorials*, vol. 17, pp. 1088–1101, Secondquarter 2015.

[2] M. R. Asghar, G. Dan, D. Miorandi, and I. Chlamtac, "Smart meter data privacy: A survey," *IEEE Communications Surveys Tutorials*, vol. 19, pp. 2820–2835, Fourthquarter 2017.

[3] L. Sankar, S. Kar, R. Tandon, and H. V. Poor, "Competitive privacy in the smart grid: An information-theoretic approach," in *IEEE International Conference on Smart Grid Communications*, pp. 220–225, Oct. 2011.

[4] D. Eckhoff and I. Wagner, "Privacy in the smart city-applications, technologies, challenges, and solutions," *IEEE Communications Surveys Tutorials*, vol. 20, pp. 489–516, Firstquarter 2018.

[5] D. He, S. Chan, and M. Guizani, "User privacy and data trustworthiness in mobile crowd sensing," *IEEE Wireless Communications*, vol. 22, pp. 28–34, Feb. 2015.

[6] R. K. Ganti, F. Ye, and H. Lei, "Mobile crowdsensing: current state and future challenges," *IEEE Communications Magazine*, vol. 49, pp. 32–39, Nov. 2011.

[7] I. J. Vergara-Laurens, L. G. Jaimes, and M. A. Labrador, "Privacy-preserving mechanisms for crowdsensing: Survey and research challenges," *IEEE Internet of Things Journal*, vol. 4, pp. 855–869, Aug. 2017.

[8] S. Gisdakis, T. Giannetsos, and P. Papadimitratos, "Security, privacy, and incentive provision for mobile crowd sensing systems," *IEEE Internet of Things Journal*, vol. 3, pp. 839–853, Oct. 2016.

[9] V. P. Crawford and J. Sobel, "Strategic information transmission," *Econometrica*, vol. 50, no. 6, pp. 1431–1451, 1982.

[10] E. Kamenica and M. Gentzkow, "Bayesian persuasion," *American Economic Review*, vol. 101, pp. 2590–2615, Oct. 2011.

[11] F. Farokhi, H. Sandberg, I. Shames, and M. Cantoni, "Quadratic Gaussian privacy games," in *IEEE Conference on Decision and Control (CDC)*, pp. 4505–4510, Dec. 2015.

[12] E. Akyol, C. Langbort, and T. Başar, "Privacy constrained information processing," in *IEEE Conference on Decision and Control (CDC)*, pp. 4511–4516, Dec. 2015.

[13] F. Farokhi and G. Nair, "Privacy-constrained communication," *IFAC-PapersOnLine*, vol. 49, pp. 43–48, Jan 2016.

[14] E. Akyol, C. Langbort, and T. Başar, "Strategic compression and transmission of information," in *IEEE Information Theory Workshop - Fall*, pp. 219–223, Oct. 2015.

[15] H. Yamamoto, "A source coding problem for sources with additional outputs to keep secret from the receiver or wiretappers (corresp.)," *IEEE Transactions on Information Theory*, vol. 29, pp. 918–923, Nov. 1983.

[16] F. P. Calmon and N. Fawaz, "Privacy against statistical inference," in *Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, pp. 1401–1408, Oct. 2012.

[17] F. P. Calmon, A. Makhdoumi, and M. Médard, "Fundamental limits of perfect privacy," in *IEEE International Symposium on Information Theory (ISIT)*, pp. 1796–1800, June 2015.

[18] C. Dwork, "Differential privacy: A survey of results," in *Theory and Applications of Models of Computation* (M. Agrawal, D. Du, Z. Duan, and A. Li, eds.), (Berlin, Heidelberg), pp. 1–19, Springer, 2008.

[19] C. Dwork and A. Roth, "The algorithmic foundations of differential privacy," *Foundations and Trends in Theoretical Computer Science*, vol. 9, no. 3–4, pp. 211–407, 2014.

[20] A. Makhdoumi, S. Salamatian, N. Fawaz, and M. Medard, "From the information bottleneck to the privacy funnel," in *IEEE Information Theory Workshop (ITW)*, pp. 501–505, Nov. 2014.

[21] L. Sankar, S. R. Rajagopalan, and H. V. Poor, "Utility-privacy tradeoffs in databases: An information-theoretic approach," *IEEE Transactions on Information Forensics and Security*, vol. 8, pp. 838–852, June 2013.

[22] B. Rassouli and D. Gündüz, "Optimal utility-privacy trade-off with total variation distance as a privacy measure," *IEEE Transactions on Information Forensics and Security*, vol. 15, pp. 594–603, 2020.

[23] S. Asoodeh, M. Diaz, F. Alajaji, and T. Linder, "Estimation efficiency under privacy constraints," *IEEE Transactions on Information Theory*, vol. 65, pp. 1512–1534, March 2019.

[24] N. Tishby, F. C. Pereira, and W. Bialek, "The information bottleneck method," in *Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, pp. 368–377, 1999.

[25] G. Chechik, A. Globerson, N. Tishby, and Y. Weiss, "Information bottleneck for Gaussian variables," *Journal of Machine Learning Research*, vol. 6, no. Jan, pp. 165–188, 2005.

[26] E. Kazıklı, S. Gezici, and S. Yüksel, "Quadratic privacy-signaling games, payoff dominant equilibria and the information bottleneck problem," *arXiv*.

[27] J. C. Harsanyi and R. Selten, *A General Theory of Equilibrium Selection in Games*. Cambridge, Massachusets: MIT Press, 1988.

[28] H. V. Poor, *An Introduction to Signal Detection and Estimation*. New York: Springer-Verlag, 1994.

[29] C. D. Aliprantis and K. C. Border, *Infinite Dimensional Analysis: A Hitchikers Guide*. Berlin: Springer-Verlag, 2006.

[30] R. Bhatia, *Matrix Analysis*. New York: Springer-Verlag, 1997.