Quadratic Privacy-Signaling Games and the MMSE Information Bottleneck Problem for Gaussian Sources

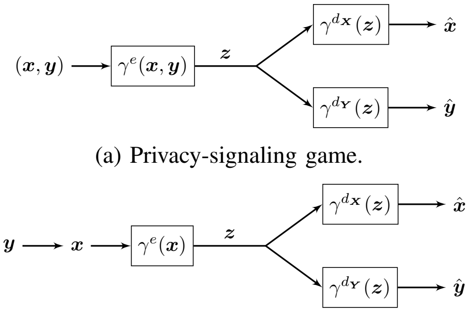
Ertan Kazıklı, Sinan Gezici, *Senior Member, IEEE,* and Serdar Yüksel, *Member, IEEE*

***Abstract*—We investigate a privacy-signaling game problem in which a sender with privacy concerns observes a pair of correlated random vectors which are modeled as jointly Gaussian. The sender aims to hide one of these random vectors and convey the other one whereas the objective of the receiver is to accurately estimate both of the random vectors. We analyze these conflicting objectives in a game theoretic framework with quadratic costs where depending on the commitment conditions (of the sender), we consider Nash or Stackelberg (Bayesian persuasion) equilibria. We show that a payoff dominant Nash equilibrium among all admissible policies is attained by a set of explicitly characterized linear policies. We also show that a payoff dominant Nash equilibrium coincides with a Stackelberg equilibrium. We formulate 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 random variables is observed at the sender. We show that this MMSE Gaussian Information Bottleneck Problem admits a linear solution which is explicitly characterized in the paper. We provide explicit conditions on when the optimal solutions, or equilibrium solutions in the Nash setup, are informative or noninformative.**

***Index Terms*—Signaling games, Nash equilibrium, Stackelberg equilibrium, privacy, estimation, information bottleneck.**

I. Introduction and System Model

Various applications such as social networks, networked control, smart grid and crowd sensing benefit from data collected from decision makers. In these applications, users share information with a service provider which wishes to improve the quality of service by utilizing information gathered from the users. The users as well are interested in enhanced service quality as they benefit from the service while at the same time they wish to retain a certain level of privacy. The privacy objective arises from the fact that the information they wish to convey to the service provider may be correlated with certain private information they want to protect. For instance, in smart grid applications, power usage information shared by the users with the service provider may disclose some information related to users such as their habits and behaviors [2], [3]. For that reason, privacy is a major challenge in smart grid applications and this is a current research topic in numerous studies (see [3]–[6] and references therein). In addition, the problem of preserving privacy while maintaining reasonable system performance appears in various contexts [7]–[22].


$(\boldsymbol{x},\boldsymbol{y})$
$\gamma^{e}(\boldsymbol{x},\boldsymbol{y})$
$\boldsymbol{z}$
$\gamma^{d_{\boldsymbol{X}}}(\boldsymbol{z})$
$\hat{\boldsymbol{x}}$
$\gamma^{d_{\boldsymbol{Y}}}(\boldsymbol{z})$
$\hat{\boldsymbol{y}}$
$\boldsymbol{y}$
$\boldsymbol{x}$
$\gamma^{e}(\boldsymbol{x})$
$\boldsymbol{z}$
$\gamma^{d_{\boldsymbol{X}}}(\boldsymbol{z})$
$\hat{\boldsymbol{x}}$
$\gamma^{d_{\boldsymbol{Y}}}(\boldsymbol{z})$
$\hat{\boldsymbol{y}}$

(a) Privacy-signaling game.

(b) Information bottleneck.

Fig. 1: Information flow under the quadratic privacy-signaling game and the information bottleneck problems.

In this paper, we consider the following communication scenario between a sender and a receiver motivated by the aforementioned applications. There is a pair of sources at the sender and the perspective of the sender is such that one source needs to be protected and the other source needs to be conveyed. As opposed to the sender, the receiver desires to accurately estimate both sources with the aim of acquiring as much information as possible. Under this setting, we investigate the interactions between the sender and the receiver whose objectives are different from each other due to the privacy concerns of the sender.

Consider an information transmission scenario in which a sender encodes a pair of correlated random vectors $\boldsymbol{X}$ and $\boldsymbol{Y}$ into $\boldsymbol{Z}$ using an encoding function denoted by $\boldsymbol{z} = \gamma^{e}(\boldsymbol{x},\boldsymbol{y})$ and a receiver wants to decode both of the random vectors based on its observation $\boldsymbol{Z} = \boldsymbol{z}$. We denote the size of the random vectors $\boldsymbol{X}$ and $\boldsymbol{Y}$ by $n_{\boldsymbol{X}}$ and $n_{\boldsymbol{Y}}$, respectively. In this communication scenario, the sender wishes to convey information contained in $\boldsymbol{Y}$ whereas it views $\boldsymbol{X}$ as a private random variable that needs to be hidden from the receiver. The aim of the receiver is to accurately estimate both of the random variables given its observation $\boldsymbol{Z} = \boldsymbol{z}$. Let the decoders for estimating $\boldsymbol{X}$ and $\boldsymbol{Y}$ at the receiver be denoted by $\gamma^{d_{\boldsymbol{X}}}(\boldsymbol{z})$ and $\gamma^{d_{\boldsymbol{Y}}}(\boldsymbol{z})$, respectively. Fig. 1a illustrates the considered information transmission scenario.

We model the random variables $\boldsymbol{X}$ and $\boldsymbol{Y}$ as jointly Gaussian random vectors. Let $\begin{bmatrix}\boldsymbol{X}\\ \boldsymbol{Y}\end{bmatrix}$ be a zero mean Gaussian

Part of this work was presented at the 2020 IEEE International Symposium on Information Theory (ISIT), June 21-26, 2020, Los Angeles, California, USA [1].

E. Kazıklı and S. Yüksel are with the Department of Mathematics and Statistics, Queen's University, K7L 3N6, Kingston, Ontario, Canada. Emails: ertan.kazikli@queensu.ca and yuksel@mast.queensu.ca.

S. Gezici is with the Department of Electrical and Electronics Engineering, Bilkent University, 06800, Ankara, Turkey, Email: gezici@ee.bilkent.edu.tr.

random vector with a positive definite covariance matrix $\Sigma \triangleq \begin{bmatrix} \Sigma_{\boldsymbol{X}} & \Sigma_{\boldsymbol{XY}} \\ \Sigma_{\boldsymbol{YX}} & \Sigma_{\boldsymbol{Y}} \end{bmatrix}$. The random variables $\boldsymbol{X}$ and $\boldsymbol{Y}$ are not independent of each other, i.e., $\Sigma_{\boldsymbol{XY}} \neq \boldsymbol{0}$. It is assumed that the joint distribution of $\boldsymbol{X}$ and $\boldsymbol{Y}$ is common knowledge, i.e., both players know $\Sigma_{\boldsymbol{X}}$, $\Sigma_{\boldsymbol{Y}}$ and $\Sigma_{\boldsymbol{XY}}$. Since $\boldsymbol{X}$ and $\boldsymbol{Y}$ are correlated, transmitting $\boldsymbol{Y}$ directly discloses information related to the private random variable $\boldsymbol{X}$. In other words, the objectives of hiding $\boldsymbol{X}$ and conveying $\boldsymbol{Y}$ are conflicting. These conflicting objectives at the sender are modeled via the following objective function:

$$\begin{aligned} &J^e(\gamma^e, \gamma^{d_{\boldsymbol{X}}}, \gamma^{d_{\boldsymbol{Y}}}) \\ &\qquad = \mathbb{E}[\|\boldsymbol{Y} - \gamma^{d_{\boldsymbol{Y}}}(\boldsymbol{Z})\|^2] - \delta\, \mathbb{E}[\|\boldsymbol{X} - \gamma^{d_{\boldsymbol{X}}}(\boldsymbol{Z})\|^2], \end{aligned} \tag{1}$$

which is to be minimized, where δ is a positive design parameter that determines the level of desired privacy in terms of hiding $\boldsymbol{X}$. On the other hand, the receiver aims to extract both of the random variables. Thus, the receiver wishes to minimize the following objective function:

$$\begin{aligned} &J^d(\gamma^e, \gamma^{d_{\boldsymbol{X}}}, \gamma^{d_{\boldsymbol{Y}}}) \\ &\qquad = \mathbb{E}[\|\boldsymbol{Y} - \gamma^{d_{\boldsymbol{Y}}}(\boldsymbol{Z})\|^2] + \mathbb{E}[\|\boldsymbol{X} - \gamma^{d_{\boldsymbol{X}}}(\boldsymbol{Z})\|^2]. \end{aligned} \tag{2}$$

In (2), the mean squared errors corresponding to random variables $\boldsymbol{X}$ and $\boldsymbol{Y}$ are incorporated into the objective function with equal weights since taking different weights does not alter the problem. In this work, we investigate the Nash equilibrium and the Stackelberg equilibrium 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: The players choose their strategies 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 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 random variable are denoted by $\Gamma^{d_{\boldsymbol{X}}}$ and $\Gamma^{d_{\boldsymbol{Y}}}$, i.e., $\gamma^{d_{\boldsymbol{X}}} \in \Gamma^{d_{\boldsymbol{X}}}$ and $\gamma^{d_{\boldsymbol{Y}}} \in \Gamma^{d_{\boldsymbol{Y}}}$. A set of policies $\gamma^{e,*}$, $\gamma^{d_{\boldsymbol{X}},*}$ and $\gamma^{d_{\boldsymbol{Y}},*}$ forms a Nash equilibrium if [23]

$$J^e(\gamma^{e,*}, \gamma^{d_{\boldsymbol{X}},*}, \gamma^{d_{\boldsymbol{Y}},*}) \leq J^e(\gamma^e, \gamma^{d_{\boldsymbol{X}},*}, \gamma^{d_{\boldsymbol{Y}},*}) \tag{3}$$

for all $\gamma^e \in \Gamma^e$ and

$$J^d(\gamma^{e,*}, \gamma^{d_{\boldsymbol{X}},*}, \gamma^{d_{\boldsymbol{Y}},*}) \leq J^d(\gamma^{e,*}, \gamma^{d_{\boldsymbol{X}}}, \gamma^{d_{\boldsymbol{Y}}}) \tag{4}$$

for all $\gamma^{d_{\boldsymbol{X}}} \in \Gamma^{d_{\boldsymbol{X}}}$ and $\gamma^{d_{\boldsymbol{Y}}} \in \Gamma^{d_{\boldsymbol{Y}}}$.

Remark 1: Under the Nash equilibrium concept, there is no commitment assumption for the players. This may be appropriate for scenarios where the players do not trust the announcements of each other or do not have access to policy announcements. For instance, a user (sender) making sensor measurements in a crowd sensing application may encounter a tradeoff between utility of providing useful information to a data aggregator (receiver) and protecting its privacy. We may consider a setting where the sender has the ability to reconfigure its policy. In this case, the sender wishes to deviate from a certain announced policy if it knows that such deviation leads to a better privacy protection given the receiver's announcement. Thus, the receiver does not trust the policy announcement of the sender. On the other hand, the sender may also think that the receiver's announcement is not trustworthy. This happens for instance when the receiver discloses collected information from individuals to third parties which do not comply with the receiver's commitment. In addition, the sender may wish to guard itself against data breaches at the legitimate receiver. In order to model such scenarios where both the sender and the receiver do not have any commitment regarding their policies, a Nash theoretic game model can be used. Although they do not commit to a certain policy, if they are in equilibrium, then they do not wish to deviate unilaterally.

On the other hand, the Stackelberg equilibrium involves a sequential game play in the sense that first the sender and then the receiver act (this setup is commonly referred to as the *Bayesian persuasion* problem in the economics literature [24]). The sender chooses and announces its strategy and then the receiver acts upon learning the strategy of the sender. Here, the sender commits to employ this announced strategy. The receiver employs an optimal response to the announced strategy of the sender. A set of policies $\gamma^{e,*}$, $\gamma^{d_{\boldsymbol{X}},*}$ and $\gamma^{d_{\boldsymbol{Y}},*}$ forms a Stackelberg equilibrium if [23]

$$\begin{aligned} &J^e(\gamma^{e,*}, \gamma^{d_{\boldsymbol{X}},*}(\gamma^{e,*}), \gamma^{d_{\boldsymbol{Y}},*}(\gamma^{e,*})) \\ &\qquad\qquad \leq J^e(\gamma^e, \gamma^{d_{\boldsymbol{X}},*}(\gamma^e), \gamma^{d_{\boldsymbol{Y}},*}(\gamma^e)) \end{aligned} \tag{5}$$

for all $\gamma^e \in \Gamma^e$, where $\gamma^{d_{\boldsymbol{X}},*}(\gamma^e)$ and $\gamma^{d_{\boldsymbol{Y}},*}(\gamma^e)$ are such that

$$\begin{aligned} &J^d(\gamma^e, \gamma^{d_{\boldsymbol{X}},*}(\gamma^e), \gamma^{d_{\boldsymbol{Y}},*}(\gamma^e)) \\ &\qquad\qquad \leq J^d(\gamma^e, \gamma^{d_{\boldsymbol{X}}}(\gamma^e), \gamma^{d_{\boldsymbol{Y}}}(\gamma^e)) \end{aligned} \tag{6}$$

for all $\gamma^{d_{\boldsymbol{X}}} \in \Gamma^{d_{\boldsymbol{X}}}$ and $\gamma^{d_{\boldsymbol{Y}}} \in \Gamma^{d_{\boldsymbol{Y}}}$.

Remark 2: It is important to emphasize that under the Stackelberg equilibrium concept, there is a commitment assumption for the sender and the sender cannot backtrack its commitment. This setup can be appropriate for scenarios where an information provider publicly shares information given its observations. For instance, consider a medical research setting. Researchers wish to publicly reveal data they obtained as a result of medical research so that other researchers can benefit from this data. However, as this data may contain sensitive information related to participants of the study, the researchers need to take privacy into account while publishing their research data. In this case, the researchers employ a privacy-preserving data revelation scheme so that the privacy of the participants is protected. On the other hand, in order for other researchers to make sense of this revealed research data, they need to know what type of privacy-preserving mechanism is employed in the design. Therefore, the researchers performing the study publicly reveal the specification of such mechanism they used. This corresponds to a scenario with sender commitment, as in the Stackelberg setup considered in this paper.

We will also consider an instance of the problem above as the *MMSE Gaussian Information Bottleneck Problem*. The difference with the setup above is that the sender only has access to $\boldsymbol{X}$, which is the message it intends to hide while revealing as much information on $\boldsymbol{Y}$ as possible. This is depicted in Fig. 1b. The classical information bottleneck problem [25] considers the mutual information as the cost criterion where the aim is to compress an observed random variable while preserving information related to an unobserved correlated random variable. Note that both the privacy and the compression objectives aim at removing the corresponding information from the revealed message. Motivated by this resemblance, we consider the information bottleneck problem in our game theoretic context. Details are provided in Section IV.

A. Literature Review

For signaling games under the Nash equilibrium concept, Crawford and Sobel in their foundational paper [26] investigate a communication scenario between a 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 transmitted message. In contrast to Crawford and Sobel, the Bayesian persuasion problem [24] investigates signaling scenarios under the Stackelberg equilibrium concept rather than the Nash equilibrium concept.

In the context of the Bayesian persuasion problem, an important related work [27] considers a multidimensional signaling scenario under a Stackelberg game setup where the sender employs a general quadratic cost structure. An upper bound on the performance of the sender is obtained via formulating a semidefinite program. For jointly Gaussian sources, a linear policy that achieves this upper bound is characterized which shows the optimality of linear policies for such a Stackelberg game. We use this characterization in some of our results rather prominently.

Recently, the strategic information transmission (SIT) problem has attracted attention also in the communication and control theory literature [28]–[38]. For instance, the work in [29] considers quadratic costs with a bias term appearing in sender's cost and investigates both scalar and multidimensional source settings. An interesting observation from [29] is the existence of a linear Nash equilibrium which is in contrast to the quantized nature of the equilibrium in Crawford and Sobel. In [30], the misalignment in the objectives is due to a bias term which is modeled as a random variable. The authors consider the Stackelberg equilibrium concept and focus on affine policies. In [31], a communication scenario between prospect theoretic agents whose cost functions are distorted by subjective biases is investigated using the Stackelberg equilibrium concept.

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 [39]–[42]. In these studies, a common theme is to model private and nonprivate random variables as jointly Gaussian random variables. In [39], 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. We investigate Nash equilibria as well and show that a payoff dominant equilibrium is attained by linear policies. We also show that these linear policies at the payoff dominant Nash equilibrium lead to a Stackelberg equilibrium even when the encoding policy is not restricted to be linear. The work in [40] also investigates a Stackelberg game where the utility measure for the nonprivate random variable is quadratic and the privacy measure is entropy based. Both noiseless and noisy communication scenarios are considered and essentially unique linear encoding and decoding policies that form a Stackelberg equilibrium are characterized. In [41], a Nash game is studied where the privacy measure is based on mutual information and the utility measure for the nonprivate random variable is quadratic. In [41], apart from the previously described Gaussian scenario, another scenario in which private and nonprivate data are treated as discrete random variables is considered.

The tradeoff between utility and privacy appears also in various other contexts [9]–[22], [43]–[47]. One line of related work is the differential privacy literature where the main problem of interest is to protect private information on publicly available databases [43], [44]. The notion of differential privacy ensures that private information provided by an individual to a database is not compromised by a third party, e.g., a data analyst, who retrieves information from this database. In this context, an interesting result from [19] is the application of the Laplacian or Gaussian perturbations to guarantee differential privacy. For a comprehensive treatment of differential privacy on such problems as filtering and estimation, please see [48]. Another line of work is the privacy funnel problem [45] where it is desired to convey as much information as possible related to an observed random variable while trying to leak as low information as possible related to an unobserved private random variable. It should be noted that in the privacy funnel problem, only the nonprivate random variable is observed at the sender whereas in our framework, we assume that both the nonprivate and private random variables are observed at the sender. Another related work is [10] where the tradeoff between utility and privacy is investigated through formulating constrained optimization problems that consider settings with a discrete random variable and a continuous random variable. The continuous random variable case focuses on Gaussian perturbations applied to the nonprivate random variable to protect private information.

As noted, a further related problem is the information bottleneck problem [25], [49]–[57] which also has connections with the privacy funnel problem [46]. In the information bottleneck technique [25], the aim is to compress an observed random variable while trying to preserve information related to another correlated random variable which is not observed. It is important to note that the information bottleneck technique is closely related to an earlier seminal work [58] which considers a similar constrained optimization problem by employing conditional entropy to asses the performance. The information bottleneck problem specializing to Gaussian sources is investigated in [59] where the random variables of interest are jointly Gaussian random vectors. 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 random variables is received at the sender whereas in our framework the costs include mean squared error terms and both of the random variables 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.

B. *Contributions of the Manuscript*

The main aim of this paper is to provide both Nash and Stackelberg equilibria analyses for the considered privacy-signaling game problem. In game theory, since the solution concept involves an equilibrium (Nash, Stackelberg, and refinements), one cannot talk about an optimal equilibrium in general. Nonetheless, as a main contribution of our work, we establish and compute an equilibrium, which is desirable among all, for both of the players. The main contributions of this paper can be summarized as follows:

- In the literature, a Nash equilibrium analysis of the privacy game problem, in which both the privacy and the utility (for the nonprivate random variable) are measured via the mean squared error cost, has not been available. In this paper, we consider this problem for the first time in the literature to our knowledge. More importantly, we show that a payoff dominant Nash equilibrium is attained by linear policies in Theorem 1. These equilibria are the most desirable equilibria for both of the players among any set of policies. We show that the characterized linear payoff dominant Nash equilibria coincide with the Stackelberg equilibria in Theorem 2. It should be emphasized that these (Stackelberg) equilibria are obtained without an *a priori* affine policy restriction for the players. In other words, if we consider the optimization problem that the encoder needs to solve while obtaining the Stackelberg equilibria, these linear policies are the optimal solution among any sets of policies.
- We introduce an MMSE Gaussian information bottleneck problem, which is a modification of the classical information bottleneck problem that has been considered under mutual information criteria. By viewing this as an instance of the privacy-signaling game under the Stackelberg formulation, we show that the solution to the MMSE Gaussian information bottleneck problem is attained by a set of explicitly characterized linear policies in Theorem 3. Namely, even when the policies are allowed to be nonlinear, a set of linear policies arises as the optimal solution.
- We extend our results for scalar sources to the additive Gaussian noise channel setting. Under this setting, it is shown that a payoff dominant Nash equilibrium is attained by linear policies in Theorem 5. This theorem also establishes that the characterized linear Nash equilibrium is unique among the affine class. We also show that the payoff dominant Nash equilibrium coincides with the Stackelberg equilibrium in Theorem 6. In addition, the characterized linear Stackelberg equilibrium is unique among any set of policies. We also establish the existence of nonlinear Nash and Stackelberg equilibria considering a discrete channel setting in which the encoding function is restricted to take discrete values in Theorem 7 and Theorem 8, respectively.

C. *Organization of the Manuscript*

The remainder of the paper is organized as follows. Section II and Section III provide, respectively, the Nash and Stackelberg equilibria analyses for the considered privacy-signaling game. Section IV investigates the information bottleneck problem as an instance of our proposed framework. Section V-A and Section V-B extend the results for scalar sources to the Gaussian noise channel and discrete channel, respectively. Section VI provides numerical examples, and Section VII concludes the paper with some final remarks.

II. Nash Equilibria

In this section, we characterize linear Nash equilibria of the considered privacy-signaling game. More importantly, we show that special cases of these equilibria lead to payoff dominant Nash equilibria. These payoff dominant Nash equilibria are the most desirable equilibria for both of the players (among all coding/decoding policies, including those that are nonlinear) in a sense that is made explicit in the following definition.

Definition 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 [60] .

In order to characterize linear Nash equilibria, we propose an equivalent formulation by applying an invertible linear transformation of variables from Tamura [27]. We note that [27] considers a general multidimensional signaling setup under quadratic costs and characterizes a set of linear policies that forms a Stackelberg equilibrium for jointly Gaussian sources. We use this characterization for our special case of privacy-signaling game scenario to formulate an equivalent problem and this approach facilitates our Nash equilibrium analysis.

Theorem 1:

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

(i) There exist informative[2] linear Nash equilibria with an encoding policy

$$\boldsymbol{z} = \begin{bmatrix}\alpha_1 \boldsymbol{q}_1 & \cdots & \alpha_{n_{\boldsymbol{Y}}} \boldsymbol{q}_{n_{\boldsymbol{Y}}}\end{bmatrix}^T \Sigma^{-1/2} \begin{bmatrix}\boldsymbol{x} \\ \boldsymbol{y}\end{bmatrix} \tag{7}$$

for scalars $\{\alpha_i\}_{i=1}^{n_{\boldsymbol{Y}}}$ with at least one of these scalars being nonzero[3] where $\{\boldsymbol{q}_i\}_{i=1}^{n}$ are normalized eigenvectors of $W = \Sigma^{1/2}\text{diag}(-\delta I, I)\Sigma^{1/2}$ with $n \triangleq n_{\boldsymbol{X}} + n_{\boldsymbol{Y}}$ and these eigenvectors are arranged in such a way that the corresponding eigenvalues $\{\lambda_i\}_{i=1}^{n}$ satisfy $\lambda_i > 0$ for $i = 1, \ldots, n_{\boldsymbol{Y}}$ and $\lambda_i < 0$ for $i = n_{\boldsymbol{Y}} + 1, \ldots, n$. The corresponding decoding policy is given by

$$\begin{bmatrix}\gamma^{d_{\boldsymbol{X}}}(\boldsymbol{z}) \\ \gamma^{d_{\boldsymbol{Y}}}(\boldsymbol{z})\end{bmatrix} = \Sigma^{1/2} \begin{bmatrix}\beta_1 \boldsymbol{q}_1 & \cdots & \beta_{n_{\boldsymbol{Y}}} \boldsymbol{q}_{n_{\boldsymbol{Y}}}\end{bmatrix} \boldsymbol{z} \tag{8}$$

where $\beta_i = 1/\alpha_i$ if $\alpha_i \neq 0$ and $\beta_i = 0$ otherwise for $i = 1, \ldots, n_{\boldsymbol{Y}}$.These Nash equilibria exist for any set of scalars $\{\alpha_i\}_{i=1}^{n_{\boldsymbol{Y}}}$. When the indices with $\alpha_i \neq 0$ are the same for two sets of scalars, they lead to the same performance values, i.e., the resulting equilibria with these sets of scalars are informationally equivalent.

(ii) These informative equilibria are payoff dominant Nash equilibria if $\alpha_i \neq 0$ for all $i = 1, \ldots, n_{\boldsymbol{Y}}$.

(iii) In addition to informative equilibria, there always exist noninformative Nash equilibria with the transmitted message being independent of the sources, e.g.,$\gamma^e(\boldsymbol{x}, \boldsymbol{y}) = C$ for some constant C, and with the decoding policies $\gamma^{d_{\boldsymbol{X}}}(\boldsymbol{z}) = \boldsymbol{0}$ and $\gamma^{d_{\boldsymbol{Y}}}(\boldsymbol{z}) = \boldsymbol{0}$.

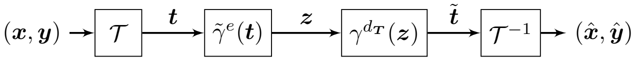
Proof: We apply a linear transformation of variables that gives an orthonormal coordinate system and then show that in this transformed coordinate system the sender wishes to convey some of the coordinates and to hide the remaining coordinates. The advantage of this transformation is that as these coordinates are orthogonal to each other, conveying one coordinate does not give information related to other coordinates.

Since $W = \Sigma^{1/2}\text{diag}(-\delta I, I)\Sigma^{1/2}$ is symmetric, we can decompose it as $W = Q\Lambda Q^T$ for orthonormal Q and diagonal Λ. We denote the columns of Q by $\{\boldsymbol{q}_i\}_{i=1}^{n}$ and the diagonal elements of Λ by $\{\lambda_i\}_{i=1}^{n}$. By Sylvester's law of inertia, W and $\text{diag}(-\delta I, I)$ have the same number of positive and negative eigenvalues [61, p. 282]. Therefore, W has $n_{\boldsymbol{Y}}$ positive and $n_{\boldsymbol{X}}$ negative eigenvalues. We sort these eigenvalues $\{\lambda_i\}_{i=1}^{n}$ in such a way that the first $n_{\boldsymbol{Y}}$ of them are positive and the remaining ones are negative. For notational convenience, we define

$$\boldsymbol{S} \triangleq \begin{bmatrix}\boldsymbol{X} \\ \boldsymbol{Y}\end{bmatrix}, \quad \gamma^{d_{\boldsymbol{S}}}(\boldsymbol{z}) \triangleq \begin{bmatrix}\gamma^{d_{\boldsymbol{X}}}(\boldsymbol{z}) \\ \gamma^{d_{\boldsymbol{Y}}}(\boldsymbol{z})\end{bmatrix}. \tag{9}$$

Now, we make the following transformation of variables:

$$\boldsymbol{T} \triangleq Q^T \Sigma^{-1/2} \boldsymbol{S}, \tag{10}$$



$(\boldsymbol{x}, \boldsymbol{y})$
$\mathcal{T}$
$\boldsymbol{t}$
$\tilde{\gamma}^e(\boldsymbol{t})$
$\boldsymbol{z}$
$\gamma^{d_{\boldsymbol{T}}}(\boldsymbol{z})$
$\tilde{\boldsymbol{t}}$
$\mathcal{T}^{-1}$
$(\hat{\boldsymbol{x}}, \hat{\boldsymbol{y}})$

Fig. 2: Equivalent formulation where $\mathcal{T}$ denotes the linear transformation specified in (10).

where $\Sigma^{-1/2}$ is well-defined since Σ is assumed to be positive definite. We note that

$$\begin{aligned}\mathbb{E}[\boldsymbol{T}\boldsymbol{T}^T] &= \mathbb{E}[Q^T \Sigma^{-1/2} \boldsymbol{S}\boldsymbol{S}^T \Sigma^{-1/2} Q] \\ &= Q^T \Sigma^{-1/2} \mathbb{E}[\boldsymbol{S}\boldsymbol{S}^T] \Sigma^{-1/2} Q \\ &= Q^T \Sigma^{-1/2} \Sigma \Sigma^{-1/2} Q = I.\end{aligned}$$

Thus, each components of $\boldsymbol{T}$ are independent and identically distributed zero-mean Gaussian random variables each with a unit variance.

Next, we propose an equivalent problem under this linear transformation of variables. The encoder consists of two consecutive mappings, one of which is fixed as above and the other one can arbitrarily be chosen by the sender. In other words, there is a linear mapping from $(\boldsymbol{x}, \boldsymbol{y})$ to $\boldsymbol{t}$ fixed as (10) and then an encoding function $\boldsymbol{z} = \tilde{\gamma}^e(\boldsymbol{t})$. At the receiver side, we also consider a similar setting. The observation at the receiver is mapped into $\tilde{\boldsymbol{t}}$ via $\gamma^{d_{\boldsymbol{T}}}(\boldsymbol{z})$, which can arbitrarily be selected by the receiver. Then, these auxiliary variables are mapped into estimates of $\boldsymbol{x}$ and $\boldsymbol{y}$ as follows:

$$\gamma^{d_{\boldsymbol{S}}}(\boldsymbol{z}) = \Sigma^{1/2} Q \gamma^{d_{\boldsymbol{T}}}(\boldsymbol{z}). \tag{11}$$

Fig. 2 provides an illustration for the equivalent formulation. The aim is to characterize $\tilde{\gamma}^e(\boldsymbol{t})$ as well as $\gamma^{d_{\boldsymbol{T}}}(\boldsymbol{z})$ at a Nash equilibrium. Since the proposed transformation is invertible, the problem in the transform domain is equivalent to the original problem.

Next, we express the objective function of each player in terms of the random variables in the transformed coordinate system. The objective function of the sender can be written as

$$\begin{aligned}J^e(\tilde{\gamma}^e, \gamma^{d_{\boldsymbol{T}}}) &= \mathbb{E}[(\boldsymbol{S} - \gamma^{d_{\boldsymbol{S}}}(\boldsymbol{Z}))^T \text{diag}(-\delta I, I) \\ &\qquad (\boldsymbol{S} - \gamma^{d_{\boldsymbol{S}}}(\boldsymbol{Z}))] \\ &= \mathbb{E}[(\boldsymbol{T} - \gamma^{d_{\boldsymbol{T}}}(\boldsymbol{Z}))^T Q^T \Sigma^{1/2} \text{diag}(-\delta I, I) \\ &\qquad \Sigma^{1/2} Q (\boldsymbol{T} - \gamma^{d_{\boldsymbol{T}}}(\boldsymbol{Z}))] \\ &= \mathbb{E}[(\boldsymbol{T} - \gamma^{d_{\boldsymbol{T}}}(\boldsymbol{Z}))^T \Lambda (\boldsymbol{T} - \gamma^{d_{\boldsymbol{T}}}(\boldsymbol{Z}))] \\ &= \sum_{i=1}^{n} \lambda_i \mathbb{E}[(T_i - \gamma^{d_{T_i}}(\boldsymbol{Z}))^2],\end{aligned} \tag{12}$$

where the first equation is obtained from (1) and (9), the second equation is based on (10) and (11), and $\{\lambda_i\}_{i=1}^{n}$ are the eigenvalues of W which satisfy $\lambda_i > 0$ for $i = 1, \ldots, n_{\boldsymbol{Y}}$ and $\lambda_i < 0$ for $i = n_{\boldsymbol{Y}} + 1, \ldots, n$. Similarly, if we express the objective function of the receiver in terms of the random variables in the introduced coordinate system, we get

$$\begin{aligned}J^d(\tilde{\gamma}^e, \gamma^{d_{\boldsymbol{T}}}) &= \mathbb{E}[(\boldsymbol{S} - \gamma^{d_{\boldsymbol{S}}}(\boldsymbol{Z}))^T (\boldsymbol{S} - \gamma^{d_{\boldsymbol{S}}}(\boldsymbol{Z}))] \\ &= \mathbb{E}[(\boldsymbol{T} - \gamma^{d_{\boldsymbol{T}}}(\boldsymbol{Z}))^T K (\boldsymbol{T} - \gamma^{d_{\boldsymbol{T}}}(\boldsymbol{Z}))]\end{aligned} \tag{13}$$

where $K \triangleq Q^T \Sigma Q$ is a positive definite matrix since Σ is positive definite. In the proposed equivalent problem, the

[2] We refer to an equilibrium as noninformative if the sender does not convey information related to both of the random variables 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.

[3] The case when $\alpha_i = 0$ for all $i = 1, \ldots, n_{\boldsymbol{Y}}$ leads to a noninformative Nash equilibrium.

optimal $\gamma^{d_{\boldsymbol{T}}}(\boldsymbol{z})$ for a given encoding policy $\tilde{\gamma}^e(\boldsymbol{t})$ is the minimum mean squared error estimator of the corresponding random variable. The proof of this statement is standard but we present it in Lemma 1 of Appendix A for completeness.

In the equivalent formulation, the objective function of the sender is expressed as a weighted sum of mean squared error terms corresponding to independent random variables where there are both positive and negative weights. We partition the transformed random vector as

$$\boldsymbol{T} \triangleq \begin{bmatrix} \boldsymbol{U} \\ \boldsymbol{V} \end{bmatrix} \tag{14}$$

where $\boldsymbol{U} \in \mathbb{R}^{n_{\boldsymbol{Y}}}$ and $\boldsymbol{V} \in \mathbb{R}^{n_{\boldsymbol{X}}}$ such that the positive and negative coefficients in (12) correspond to $\boldsymbol{U}$ and $\boldsymbol{V}$, respectively. The receiver still employs the minimum mean squared error estimators for the corresponding random variables for a given encoding policy. Next, we use the equivalent formulation to characterize the Nash equilibria. Due to structure of sender's cost in (12) considering the equivalent formulation, it follows that the sender does not convey any information related to $\boldsymbol{V}$, which is specified in (14), at a Nash equilibrium. We present this auxiliary result in Lemma 2 of Appendix A. Hence, the transmitter is restricted to send information related to $\boldsymbol{U}$ at a Nash equilibrium. This implies that the objective function of the sender at a Nash equilibrium reduces to

$$J^e(\tilde{\gamma}^e, \gamma^{d_{\boldsymbol{T}}}) = \sum_{i=1}^{n_{\boldsymbol{Y}}} \lambda_i \mathbb{E}[(T_i - \gamma^{d_{T_i}}(\boldsymbol{Z}))^2] \tag{15}$$

where $\{1, \ldots, n_{\boldsymbol{Y}}\}$ are the only indices with $\lambda_i > 0$ in $\{\lambda_i\}_{i=1}^n$. Since the receiver wishes to extract all the random variables in this transformed coordinate system including the ones with these indices, the receiver also shares the objective of minimizing these mean squared error terms. As a result, conveying all or a subset of the random variables in $\boldsymbol{U}$ yields a Nash equilibrium and this gives the linear policies stated in (7). In addition, conveying all of these random variables yields the minimum attainable performance at a Nash equilibrium for both of the players. Thus, if $\alpha_i \neq 0$ for all $i = 1, \ldots, n_{\boldsymbol{Y}}$ in (7), the corresponding linear Nash equilibria are payoff dominant Nash equilibria. Namely, there does not exist any other Nash equilibrium which Pareto dominates these characterized equilibria. ■

Remark 3: It is interesting to contrast the result of Theorem 1 with the signaling game setups in the literature where the misaligned cost structure arises from biased nature of the sender as opposed to privacy concerns of the sender. Notably, Crawford and Sobel [26] introduce a signaling game setup where the costs are misaligned due to a bias term. The communication setup is similar to our setting in this section in the sense that the transmitted message is perfectly observed by the receiver and the sender does not have a power constraint. Under this setting, [26] establishes the quantized nature of Nash equilibria for scalar sources supported on $[0, 1]$ under certain assumptions regarding the objectives, and [29] generalizes this result to arbitrary source distributions under quadratic criteria. In contrast to this setting with a biased sender, if there is a privacy concerned sender, then a Nash equilibrium is attained by linear policies regardless of source dimensions, as shown in Theorem 1. In fact, depending on whether the source is scalar or vector valued, there may exist linear informative Nash equilibria for the signaling game setup with a biased sender. In particular, in [38], we extend Crawford and Sobel's formulation to a multidimensional source setting under quadratic cost criteria with a biased sender and show that for independent and identically distributed sources and an arbitrary bias vector, there always exist linear informative Nash equilibria (only) when the source distribution is Gaussian.

Theorem 1 characterizes linear Nash equilibria in which there is communication between the transmitter and the receiver. Hence, the considered game always admits informative linear Nash equilibria regardless of the system parameters. More importantly, special cases of these linear equilibria leads to payoff dominant Nash equilibria, which are the most desirable equilibria for both of the players among all coding/decoding policies.

Next, we specialize to the case of scalar sources with the aim to provide a more explicit characterization of the payoff dominant Nash equilibria in this case. In particular, X and Y are assumed to be zero-mean jointly Gaussian with variances σ_X^2 and σ_Y^2, respectively and a nonzero correlation ρ, i.e., $\Sigma = \begin{bmatrix} \sigma_X^2 & \rho \\ \rho & \sigma_Y^2 \end{bmatrix}$. We present the following as a corollary of Theorem 1.

Corollary 1: For scalar sources X and Y, there exist informative linear Nash equilibria with an encoding policy $\gamma^e(x, y) = Ax + By$ which satisfies

$$\frac{B}{A} = -\frac{\delta\sigma_X^2 + \sigma_Y^2}{2\delta\rho} - \frac{\sqrt{(\delta\sigma_X^2 + \sigma_Y^2)^2 - 4\delta\rho^2}}{2\delta\rho} \tag{16}$$

and decoding policies

$$\gamma^{d_X}(z) = \left(\frac{A\sigma_X^2 + B\rho}{A^2\sigma_X^2 + B^2\sigma_Y^2 + 2AB\rho}\right) z, \tag{17}$$

$$\gamma^{d_Y}(z) = \left(\frac{A\rho + B\sigma_Y^2}{A^2\sigma_X^2 + B^2\sigma_Y^2 + 2AB\rho}\right) z. \tag{18}$$

These equilibria are payoff dominant Nash equilibria. In addition, these equilibria are unique among linear policies.

Proof: In order to characterize the Nash equilibria, we need to find the eigenvalues and eigenvectors of $W = \Sigma^{1/2}\text{diag}(-\delta, 1)\Sigma^{1/2}$. We note that W has the same eigenvalues as $D \triangleq \Sigma\,\text{diag}(-\delta, 1)$ and these can be computed as

$$\lambda_1 = \frac{-\delta\sigma_X^2 + \sigma_Y^2}{2} + \frac{\sqrt{(\delta\sigma_X^2 + \sigma_Y^2)^2 - 4\delta\rho^2}}{2}, \tag{19}$$

$$\lambda_2 = \frac{-\delta\sigma_X^2 + \sigma_Y^2}{2} - \frac{\sqrt{(\delta\sigma_X^2 + \sigma_Y^2)^2 - 4\delta\rho^2}}{2}. \tag{20}$$

where $\lambda_1 > 0$ and $\lambda_2 < 0$. Thus, the sender is restricted to transmit $u = \boldsymbol{q}_1^T\Sigma^{-1/2}\begin{bmatrix} x \\ y \end{bmatrix}$ where $\boldsymbol{q}_1$ is the normalized eigenvector of W associated with the eigenvalue λ_1 computed in (19). It is seen that $\boldsymbol{q}_1^T\Sigma^{-1/2}$ is a left eigenvector of D associated with its eigenvalue λ_1. By expressing this left eigenvector, an encoding policy which satisfies (16) is obtained. Then, the conditional expectation formula for jointly Gaussian distributions can be employed to obtain (17) and (18) [62, p. 155]. As a result, these characterized policies lead to payoff dominant Nash equilibria. Moreover, the only possible linear equilibria are attained by transmitting a scaled version of u. ■

Remark 4: Note that when $\delta \to 0^+$, the encoding policy satisfies $A/B \to 0$ as can be deduced from (16). Therefore, $\delta \to 0^+$ implies that the encoder transmits directly Y at a Nash equilibrium. When $\delta = 0$, the sender also transmits Y directly at a Nash equilibrium since it does not have any privacy concern in this case. Hence, the Nash equilibrium specified in Corollary 1 as $\delta \to 0^+$ coincides with the Nash equilibrium when $\delta = 0$.

Remark 5: It is seen that the ratio of B and A converges to $-\sigma_x^2/\rho$ as $\delta \to \infty$, which can be verified from (16). This shows that in the high privacy regime, the encoder makes the revealed information Z and the private random variable X essentially uncorrelated.

III. Stackelberg Equilibria

In this section, we characterize the Stackelberg equilibria of the considered quadratic privacy-signaling game. Our main result is that the payoff dominant Nash equilibria characterized in the previous section are also Stackelberg equilibria. It is important to note that the set of possible encoding strategies, i.e., Γ^e, is not restricted to be linear. If the sender performs an optimization of its objective function by anticipating the best response of the receiver, these linear policies become the optimal solution among any set of policies.

We note that the Stackelberg equilibria can be obtained from the analysis presented by Tamura [27, Theorem 2] which characterizes linear policies that form a Stackelberg equilibrium for a general quadratic multidimensional signaling setup for jointly Gaussian sources with some generalizations (in the cost setup considered) but also some restrictions, such as the *a priori* restriction of the decoder to an affine class in the conditional expectation (that limits the applicability for the noisy channel setup to be considered later in the paper.) Accordingly, we consider an alternative approach where we use the equivalent formulation employed in Section II.

Theorem 2:

(i) The encoding policies in (7) with $\alpha_i \neq 0$ for all $i = 1, \dots, n_{\boldsymbol{Y}}$ and the decoding policies in (8) form a Stackelberg equilibrium, i.e., a payoff dominant Nash equilibrium and a Stackelberg equilibrium coincide.

(ii) In contrast to Nash setup for which there exist both informative and noninformative equilibria, a Stackelberg equilibrium is always informative, where the sender uses the private and nonprivate random variables in constructing its message.

(iii) When the nonprivate random variable is not scalar valued, i.e., $n_{\boldsymbol{Y}} > 1$, there exist informative Nash equilibria which do not coincide with the Stackelberg equilibria. These Nash equilibria are attained by an encoding policy in (7) and a decoding policy in (8) where the scalars $\{\alpha_i\}_{i=1}^{n_{\boldsymbol{Y}}}$ can take any value with at least one zero term and one nonzero term.

Proof: We apply a transformation of variables as in (10). In this transformed coordinate system, the objective functions of the sender and the receiver are expressed as in (12) and (13), respectively. If we partition the random vector $\boldsymbol{T}$ in this coordinate system as in (14) according to the sign of coefficients in (12), we can show that the sender can only convey information related to $\boldsymbol{U}$. In particular, in Lemma 3 of Appendix A, we prove that the sender does not convey information related to $\boldsymbol{V}$ at a Stackelberg equilibrium, which is proven in a similar manner to the proof of Lemma 2 with the exception that in this case the sender announces its policy first.

As a result of Lemma 3, the sender is restricted to transmit $\boldsymbol{U}$. Since the linear encoding policies in (7) with $\alpha_i \neq 0$ for all $i = 1, \dots, n_{\boldsymbol{Y}}$ reveals $\boldsymbol{U}$ completely, these encoding policies yield the minimum attainable performance for the sender among all encoding policies. Therefore, the policies in the statement of the theorem lead to a Stackelberg equilibrium.

We note that the random variable $\boldsymbol{U}$, which is desired to be conveyed at a Stackelberg equilibrium, has a size of $n_{\boldsymbol{Y}} \geq 1$. Therefore, there always exists informative Stackelberg equilibrium where the sender reveals $\boldsymbol{U}$ completely. On the other hand, it is always possible to construct a noninformative Nash equilibrium as noted in Theorem 1.

When $n_{\boldsymbol{Y}} > 1$, the random variable $\boldsymbol{U}$ is not scalar valued. Thus, there exist informative Nash equilibria with an encoding policy in (7) where the scalars $\{\alpha_i\}_{i=1}^{n_{\boldsymbol{Y}}}$ take any value with at least one zero term and one nonzero term. In these informative Nash equilibria, the sender conveys only a subset of the random variables in $\boldsymbol{U}$. Since at these informative Nash equilibria, the performance of the sender is strictly worse than that at the payoff dominant Nash equilibria, these policies do not lead to a Stackelberg equilibrium. ■

Remark 6: As noted earlier, [39] considers a Stackelberg game setup where there is also side information at the receiver. In particular, [39, Theorem 3] makes an *a priori* affine policy restriction and provides an implicit characterization for the equilibrium solution in terms of an optimization problem. On the other hand, Theorem 2 of our paper does not make an *a priori* affine policy restriction and provides an explicit characterization of the Stackelberg equilibrium solution. In that sense, our result reveals that linear policies arise as the equilibrium solution among any set of policies for the Stackelberg game setup without receiver side information. In addition, if we consider the optimization problem that specifies equilibrium policies under the affine policy restriction in [39, Theorem 3] and ignore the receiver side information, then the policies in Theorem 2 of our manuscript give the optimal solution, as expected.

Remark 7: It is interesting to contrast our results in the case of a privacy concerned sender with the results in the strategic information transmission literature involving a biased sender. If one considers the classical setup of Crawford and Sobel [26] with a biased sender under the Stackelberg equilibrium concept (rather than the Nash equilibrium concept as investigated in [26]), then there exist linear equilibria [29]. In addition, [28] considers a Gaussian signaling game setup with a biased sender where the bias is modeled as a random variable and shows the optimality of linear policies for the scalar case. Hence, similar to our setup with a privacy concerned sender, in the setups with a biased sender, the Stackelberg equilibrium solutions are given by linear policies in the case of Gaussian sources.

Remark 8: In fact, employing any invertible function $h(\boldsymbol{u})$ at the sender also yields a Stackelberg equilibrium. Since the receiver knows the commitment of the sender, it can simply employ $h^{-1}(\boldsymbol{u})$ to perfectly recover $\boldsymbol{u}$.

IV. The MMSE Gaussian Information Bottleneck Problem

We now visit and re-formulate the information bottleneck problem [25] as depicted in Figure 1b. This problem has gained a significant attention in the recent literature [46], [49]–[57]. We will interpret this problem as an instance of our formulation under the Stackelberg equilibrium concept in the following sense: in contrast to the privacy game setup, only the random variable $\boldsymbol{X}$ is observed at the sender in the information bottleneck setup. In particular, we provide an estimation theoretic perspective on the information bottleneck problem by using quadratic distortion criteria as in our privacy game setup.

The information bottleneck problem considers a similar objective to that employed in this paper where the performance metrics involve mutual information rather than the mean squared error. In the information bottleneck problem, the aim is to compress an observed random variable while trying to preserve information related to a correlated random variable. These conflicting objectives are analyzed by formulating an optimization problem involving the mutual information between the considered random variables. Although the problem can be cast as a constrained optimization problem of maximizing the released (useful) information related to the unobserved random variable under a compression constraint, the Lagrangian dual of this constrained problem is commonly considered (see Remark 11 for a constrained version in our setting). The aim is to find the optimal solution to the following optimization problem:

$$\min_{\boldsymbol{Z}=\gamma^e(\boldsymbol{X})} I(\boldsymbol{X};\boldsymbol{Z}) - \beta I(\boldsymbol{Y};\boldsymbol{Z}), \tag{21}$$

where $\beta \geq 0$ is a tradeoff parameter. Here, the goal is to compress an observation $\boldsymbol{X}$ while at the same time to maximize the released information related to $\boldsymbol{Y}$.

The information bottleneck problem considering jointly Gaussian multidimensional sources is studied in [59] where the structure of the optimal solution, which is jointly Gaussian with $\boldsymbol{X}$ [63], is identified. The objective function in (21) resembles the objective function considered in this paper in the sense that in both problems the random variable $\boldsymbol{Y}$ is desired to be conveyed while information related to the random variable $\boldsymbol{X}$ is desired to be removed from the displayed message.

The information bottleneck problem can in fact be viewed as a Stackelberg game between a sender and a receiver. In this game, the sender wants to compress the observed random variable and to convey the unobserved random variable. The use of mutual information as a performance metric effectively means that the receiver uses all the available information related to both of the random variables, i.e., it always employs its best response as in the Stackelberg equilibrium. Thus, the receiver is concerned with extracting information related to both of the random variables, which is also the case in our framework.

In the following, we consider a setting which is similar to the information bottleneck technique by employing mean squared error terms as our metric. As in the information bottleneck framework, the sender observes only the random variable $\boldsymbol{X}$, rather than observing both of the random variables. Namely, the encoder has access to only partial information and is of the form $\boldsymbol{z} = \gamma^e(\boldsymbol{x})$. The objective functions of the sender and receiver are as defined in (1) and (2), respectively. Since the receiver is concerned with estimating both of the random variables, it employs the minimum mean squared error estimators of each random variable. Since the equilibrium concept is the Stackelberg equilibrium, the objective function of the sender can be written as

$$J^e(\gamma^e) = -\delta\,\mathbb{E}[\|\boldsymbol{X}-\mathbb{E}[\boldsymbol{X}|\boldsymbol{Z}]\|^2] + \mathbb{E}[\|\boldsymbol{Y}-\mathbb{E}[\boldsymbol{Y}|\boldsymbol{Z}]\|^2]. \tag{22}$$

We now present the MMSE Gaussian information bottleneck solution.

Theorem 3:

(i) When $\Sigma_{\boldsymbol{XY}}\Sigma_{\boldsymbol{YX}} - \delta\Sigma_{\boldsymbol{X}}^2$ is not negative definite, the MMSE Gaussian information bottleneck solution is attained by an encoding policy

$$\boldsymbol{z} = \begin{bmatrix}\alpha_1\boldsymbol{q}_1 & \ldots & \alpha_k\boldsymbol{q}_k\end{bmatrix}^T \Sigma_{\boldsymbol{X}}^{-1/2}\boldsymbol{x} \tag{23}$$

for nonzero scalars $\{\alpha_i\}_{i=1}^k$ where k denotes the number of nonnegative eigenvalues of $W = \Sigma_{\boldsymbol{X}}^{-1/2}\Sigma_{\boldsymbol{XY}}\Sigma_{\boldsymbol{YX}}\Sigma_{\boldsymbol{X}}^{-1/2} - \delta\Sigma_{\boldsymbol{X}}$ and $\{\boldsymbol{q}_i\}_{i=1}^k$ are the normalized eigenvectors of W associated with its nonnegative eigenvalues. The corresponding decoding policy is given by

$$\begin{bmatrix}\gamma^{d_{\boldsymbol{X}}}(\boldsymbol{z})\\ \gamma^{d_{\boldsymbol{Y}}}(\boldsymbol{z})\end{bmatrix} = \mathrm{diag}(\Sigma_{\boldsymbol{X}}^{1/2}, \Sigma_{\boldsymbol{YX}}\Sigma_{\boldsymbol{X}}^{-1/2}) \times \begin{bmatrix}\beta_1\boldsymbol{q}_1 & \cdots & \beta_k\boldsymbol{q}_k\end{bmatrix}\boldsymbol{z} \tag{24}$$

where $\beta_i = 1/\alpha_i$ for $i = 1,\ldots,k$.

(ii) In the particular case when $\Sigma_{\boldsymbol{XY}}\Sigma_{\boldsymbol{YX}} - \delta\Sigma_{\boldsymbol{X}}^2$ is positive definite, the MMSE Gaussian information bottleneck solution is attained by a fully informative encoding policy, where the sender reveals the random variable $\boldsymbol{X}$ completely.

(iii) When $\Sigma_{\boldsymbol{XY}}\Sigma_{\boldsymbol{YX}} - \delta\Sigma_{\boldsymbol{X}}^2$ is negative definite, the MMSE Gaussian information bottleneck solution is noninformative, i.e., the sender does not reveal any information related to its observation.

Before we present the proof, we contrast our estimation theoretic solution of information bottleneck problem in Theorem 3 with the information theoretic solution in [59]. Towards that goal, we restate [59, Theorem 3.1] which gives the solution to the optimization problem in (21).

Theorem 4 ([59, Theorem 3.1]): For the Gaussian information bottleneck problem under (21), the optimal solution is given by

$$\boldsymbol{z} = A(\beta)\boldsymbol{x} + \boldsymbol{n} \tag{25}$$

where $\boldsymbol{n}$ is a realization of a zero-mean Gaussian random vector with identity covariance matrix and

$$A(\beta) = \begin{cases} [\mathbf{0} \ \cdots \ \mathbf{0}]^T & \text{if } 0 \leq \beta < \beta_1^c \\ [\alpha_1 \boldsymbol{p}_1 \ \mathbf{0} \ \cdots \ \mathbf{0}]^T & \text{if } \beta_1^c \leq \beta < \beta_2^c \\ [\alpha_1 \boldsymbol{p}_1 \ \alpha_2 \boldsymbol{p}_2 \ \mathbf{0} \ \cdots \ \mathbf{0}]^T & \text{if } \beta_2^c \leq \beta < \beta_3^c \\ \ \vdots & \end{cases}$$

where $\{\boldsymbol{p}_i^T\}_{i=1}^{n_{\boldsymbol{X}}}$ are left eigenvectors of

$$\Sigma_{\boldsymbol{X}|\boldsymbol{Y}} \Sigma_{\boldsymbol{X}}^{-1} \triangleq (\Sigma_{\boldsymbol{X}} - \Sigma_{\boldsymbol{X}\boldsymbol{Y}} \Sigma_{\boldsymbol{Y}}^{-1} \Sigma_{\boldsymbol{Y}\boldsymbol{X}}) \Sigma_{\boldsymbol{X}}^{-1}$$

with the corresponding eigenvalues $\{\lambda_i\}_{i=1}^{n_{\boldsymbol{X}}}$, which are sorted in ascending order, $\beta_i^c = 1/(1-\lambda_i)$ and $\alpha_i = \sqrt{\frac{\beta(1-\lambda_i)-1}{\lambda_i(\boldsymbol{p}_i^T \Sigma_{\boldsymbol{X}} \boldsymbol{p}_i)}}$ for $i = 1, \ldots, n_{\boldsymbol{X}}$.

Remark 9: We note that for the information bottleneck problem under (21), the optimal solution is jointly Gaussian with $\boldsymbol{X}$ which is also the case for our estimation theoretic solution in Theorem 3. However, in the information theoretic formulation, the solution involves applying an independent Gaussian perturbation to a linear function of $\boldsymbol{X}$ as given in (25) whereas in our estimation theoretic solution the encoder reveals a linear function of $\boldsymbol{X}$ without applying any perturbation. We note that if one considers the original unconstrained version of the information bottleneck problem, then randomization will be needed to ensure that the constraint condition holds with an equality in some constraint regime. See Remark 11 for further details.

Next, we present the proof of Theorem 3.

Proof: We have that $(\boldsymbol{Y} - \mathbb{E}[\boldsymbol{Y}|\boldsymbol{X}])$ is orthogonal to $\boldsymbol{X}$ since

$$\mathbb{E}[\boldsymbol{Y}\boldsymbol{X}] = \mathbb{E}[\mathbb{E}[\boldsymbol{Y}\boldsymbol{X}|\boldsymbol{X}]] = \mathbb{E}[\mathbb{E}[\boldsymbol{Y}|\boldsymbol{X}]\boldsymbol{X}].$$

Since the sender observes only the random variable $\boldsymbol{X}$ and determines its message based on $\boldsymbol{X}$, it follows that $\boldsymbol{Y} \rightarrow \boldsymbol{X} \rightarrow \boldsymbol{Z}$ is a Markov chain in that order. This Markov property implies that $(\boldsymbol{Y} - \mathbb{E}[\boldsymbol{Y}|\boldsymbol{X}])$ is orthogonal to $\boldsymbol{Z}$ as well. To see this, observe that

$$\mathbb{E}[\boldsymbol{Z}\mathbb{E}[\boldsymbol{Y}|\boldsymbol{X}]] = \mathbb{E}[\boldsymbol{Z}\mathbb{E}[\boldsymbol{Y}|\boldsymbol{X},\boldsymbol{Z}]] = \mathbb{E}[\mathbb{E}[\boldsymbol{Z}\boldsymbol{Y}|\boldsymbol{X},\boldsymbol{Z}]] = \mathbb{E}[\boldsymbol{Z}\boldsymbol{Y}]$$

where the first equality is due to the Markov property and the last equality is due to iterated expectations. By using these orthogonality properties, we can express the second term in (22) as

$$\begin{aligned} &\mathbb{E}[\|\boldsymbol{Y} - \mathbb{E}[\boldsymbol{Y}|\boldsymbol{Z}]\|^2] \\ &= \mathbb{E}[\|\boldsymbol{Y} - \mathbb{E}[\boldsymbol{Y}|\boldsymbol{X}] + \mathbb{E}[\boldsymbol{Y}|\boldsymbol{X}] - \mathbb{E}[\boldsymbol{Y}|\boldsymbol{Z}]\|^2] \\ &= \mathbb{E}[\|\boldsymbol{Y} - \mathbb{E}[\boldsymbol{Y}|\boldsymbol{X}]\|^2] + \mathbb{E}[\|\mathbb{E}[\boldsymbol{Y}|\boldsymbol{X}] - \mathbb{E}[\boldsymbol{Y}|\boldsymbol{Z}]\|^2] \\ &= \mathbb{E}[\|\boldsymbol{Y} - \mathbb{E}[\boldsymbol{Y}|\boldsymbol{X}]\|^2] + \mathbb{E}[\|\mathbb{E}[\boldsymbol{Y}|\boldsymbol{X}] - \mathbb{E}[\mathbb{E}[\boldsymbol{Y}|\boldsymbol{X}]|\boldsymbol{Z}]\|^2] \end{aligned}$$

where the second equality follows from the fact that $(\boldsymbol{Y} - \mathbb{E}[\boldsymbol{Y}|\boldsymbol{X}])$ is orthogonal to $\boldsymbol{X}$ and $\boldsymbol{Z}$ and the third equality is due to iterated expectations. Observing that

$$\mathbb{E}[\boldsymbol{Y}|\boldsymbol{X} = \boldsymbol{x}] = \Sigma_{\boldsymbol{Y}\boldsymbol{X}} \Sigma_{\boldsymbol{X}}^{-1} \boldsymbol{x},$$

the objective function can be written as

$$\begin{aligned} &J^e(\gamma^e) = \mathbb{E}[\|\boldsymbol{Y} - \mathbb{E}[\boldsymbol{Y}|\boldsymbol{X}]\|^2] - \delta\,\mathbb{E}[\|\boldsymbol{X} - \mathbb{E}[\boldsymbol{X}|\boldsymbol{Z}]\|^2] \\ &+ \mathbb{E}[(\boldsymbol{X} - \mathbb{E}[\boldsymbol{X}|\boldsymbol{Z}])^T (\Sigma_{\boldsymbol{X}}^{-1} \Sigma_{\boldsymbol{X}\boldsymbol{Y}} \Sigma_{\boldsymbol{Y}\boldsymbol{X}} \Sigma_{\boldsymbol{X}}^{-1})(\boldsymbol{X} - \mathbb{E}[\boldsymbol{X}|\boldsymbol{Z}])], \end{aligned}$$

where the first term is independent of the encoder. Thus, we obtain an optimization problem of the form

$$\min_{\boldsymbol{Z}=\gamma^e(\boldsymbol{X})} \mathbb{E}\Big[(\boldsymbol{X} - \mathbb{E}[\boldsymbol{X}|\boldsymbol{Z}])^T M (\boldsymbol{X} - \mathbb{E}[\boldsymbol{X}|\boldsymbol{Z}])\Big], \tag{26}$$

where

$$M \triangleq (\Sigma_{\boldsymbol{X}}^{-1} \Sigma_{\boldsymbol{X}\boldsymbol{Y}} \Sigma_{\boldsymbol{Y}\boldsymbol{X}} \Sigma_{\boldsymbol{X}}^{-1} - \delta I).$$

The optimization problem in (26) can be viewed as a quadratic multidimensional signaling game problem considered earlier in the paper and the solution can be obtained by using the analysis in Section III. In particular, we can rewrite the problem in (26) as a Stackelberg game between a sender and a receiver with objective functions

$$J^e(\gamma^e, \gamma^{d_{\boldsymbol{X}}}) = \mathbb{E}[(\boldsymbol{X} - \gamma^{d_{\boldsymbol{X}}}(\boldsymbol{Z}))^T M (\boldsymbol{X} - \gamma^{d_{\boldsymbol{X}}}(\boldsymbol{Z}))], \tag{27}$$

$$J^d(\gamma^e, \gamma^{d_{\boldsymbol{X}}}) = \mathbb{E}[(\boldsymbol{X} - \gamma^{d_{\boldsymbol{X}}}(\boldsymbol{Z}))^T (\boldsymbol{X} - \gamma^{d_{\boldsymbol{X}}}(\boldsymbol{Z}))]. \tag{28}$$

Notice that for a given encoding policy, the best response of the receiver under (28) is given by the minimum mean squared estimator of $\boldsymbol{X}$ given $\boldsymbol{Z}$, which is consistent with (26).

Next, we apply a transformation of variables to express the objective function of the sender in a simplified form. Towards that goal, let $W = \Sigma_{\boldsymbol{X}}^{-1/2} \Sigma_{\boldsymbol{X}\boldsymbol{Y}} \Sigma_{\boldsymbol{Y}\boldsymbol{X}} \Sigma_{\boldsymbol{X}}^{-1/2} - \delta \Sigma_{\boldsymbol{X}} = Q \Lambda Q^T$ for orthonormal Q and diagonal Λ. Now, consider the following invertible transformation of variables

$$\boldsymbol{T} \triangleq Q^T \Sigma_{\boldsymbol{X}}^{-1/2} \boldsymbol{X}. \tag{29}$$

Under this transformation of variables, we introduce an equivalent problem in a similar manner to Section II. The observation $\boldsymbol{X}$ is mapped into $\boldsymbol{T}$ via a fixed transformation (29). The encoder chooses an arbitrary policy $\tilde{\gamma}^e(\cdot)$ which maps the transformed random vector $\boldsymbol{T}$ into the message $\boldsymbol{Z}$. The receiver applies an arbitrary policy $\gamma^{d_{\boldsymbol{T}}}(\cdot)$ to its observation $\boldsymbol{Z}$ and then the estimate of $\boldsymbol{X}$ is obtained via a fixed relation

$$\gamma^{d_{\boldsymbol{X}}}(\boldsymbol{z}) = \Sigma_{\boldsymbol{X}}^{1/2} Q \gamma^{d_{\boldsymbol{T}}}(\boldsymbol{z}). \tag{30}$$

In this transformed coordinate system, for a given encoding policy the receiver still employs the minimum mean squared estimator of the random variable $\boldsymbol{T}$, which can be established via a similar analysis to that in Lemma 1. The objective function of the sender in this transformed coordinate system can be written

$$J^e(\tilde{\gamma}^e, \gamma^{d_{\boldsymbol{T}}}) = \sum_{i=1}^{n_{\boldsymbol{X}}} \lambda_i \mathbb{E}[(T_i - \gamma^{d_{T_i}}(\boldsymbol{Z}))^2], \tag{31}$$

where $\{\lambda_i\}_{i=1}^{n_{\boldsymbol{X}}}$ are eigenvalues of W.

If all of these eigenvalues $\{\lambda_i\}_{i=1}^{n_{\boldsymbol{X}}}$ are positive, which is equivalent to $\Sigma_{\boldsymbol{X}\boldsymbol{Y}} \Sigma_{\boldsymbol{Y}\boldsymbol{X}} - \delta \Sigma_{\boldsymbol{X}}^2$ being positive definite, then the minimum can be attained by revealing $\boldsymbol{T}$, which corresponds to the fully informative scenario. In case all of these eigenvalues are negative, i.e., $\Sigma_{\boldsymbol{X}\boldsymbol{Y}} \Sigma_{\boldsymbol{Y}\boldsymbol{X}} - \delta \Sigma_{\boldsymbol{X}}^2$ is negative definite, then revealing information related to any component of $\boldsymbol{T}$ is not desirable for the sender, and thus, this scenario leads to a noninformative Stackelberg equilibrium. In the remaining case, i.e., $\Sigma_{\boldsymbol{XY}}\Sigma_{\boldsymbol{YX}} - \delta\Sigma_{\boldsymbol{X}}^2$ is neither positive definite nor negative definite, we partition the transformed vector according to the sign of the coefficients $\{\lambda_i\}_{i=1}^{n_{\boldsymbol{X}}}$ in (31) as follows:

$$\boldsymbol{T} \triangleq \begin{bmatrix} \boldsymbol{U} \\ \boldsymbol{V} \end{bmatrix} \tag{32}$$

where $\boldsymbol{U} \in \mathbb{R}^k$ and $\boldsymbol{V} \in \mathbb{R}^{n_{\boldsymbol{X}}-k}$ correspond to nonnegative and negative coefficients in (31), respectively, with k denoting the number of nonnegative eigenvalues of W. Next, we can apply Lemma 3 for this particular Stackelberg game setup to establish that the sender cannot convey information related to $\boldsymbol{V}$ and is restricted to send information related to $\boldsymbol{U}$. As the encoding policy in (23) reveals $\boldsymbol{U}$ completely, this encoding policy achieves the minimum attainable for the sender among any set of policies. Thus, the pair of policies (23) and (24) yield a Stackelberg equilibrium, which gives the solution to the MMSE Gaussian information bottleneck problem. ■

Remark 10: As indicated in Theorem 3, in the information bottleneck setup, the solution may be informative or noninformative depending on the tradeoff parameter δ. In contrast, in the privacy-signaling setup investigated in Section III, the equilibrium solution is always informative regardless of δ as shown in Theorem 2. We can have the following interpretation regarding these results: In the privacy-signaling setup, the sender can perform perfect removal of information in the revealed message according to its objective as it has access to both of the random variables. It turns out that this is attained via a linear encoding policy for the case of Gaussian sources. On the other hand, in the information bottleneck setup, the sender having access to partial information cannot apply perfect information removal. Instead, the sender does what is best given the partial information that it has. This happens to be a full disclosure, a partial disclosure, or a no disclosure policy depending on δ.

In the special case of scalar sources, Theorem 3 simplifies. In particular, depending on the value of δ, the equilibrium is either fully informative or noninformative and we summarize this result in the following corollary.

Corollary 2: The Stackelberg equilibrium of the information bottleneck problem for scalar sources is given by one of the following cases:

(i) If $(\rho^2/\sigma_X^4) > \delta$, then the sender completely reveals X.
(ii) If $(\rho^2/\sigma_X^4) < \delta$, then the sender does not reveal information related to X.
(iii) If $(\rho^2/\sigma_X^4) = \delta$, then both informative and noninformative scenarios lead to a Stackelberg equilibrium.

Remark 11: [**Constrained MMSE Information Bottleneck Problem.**] It is also possible to apply the ideas used in the proof of Theorem 3 to a constrained version of the MMSE Gaussian information bottleneck problem where the aim is to minimize the mean squared error for estimating $\boldsymbol{Y}$ under constraint that the mean squared error for estimating $\boldsymbol{X}$ is above a certain threshold α. In this case, the problem is defined with

$$\begin{aligned} &\min_{\boldsymbol{Z}=\gamma^e(\boldsymbol{X})} \operatorname{Tr}\Big(\Upsilon\,\mathbb{E}[(\boldsymbol{X}-\mathbb{E}[\boldsymbol{X}|\boldsymbol{Z}])(\boldsymbol{X}-\mathbb{E}[\boldsymbol{X}|\boldsymbol{Z}])^T]\Big) \\ &\text{subject to}\operatorname{Tr}\Big(\mathbb{E}[(\boldsymbol{X}-\mathbb{E}[\boldsymbol{X}|\boldsymbol{Z}])(\boldsymbol{X}-\mathbb{E}[\boldsymbol{X}|\boldsymbol{Z}])^T]\Big) \geq \alpha \end{aligned} \tag{33}$$

where $\Upsilon \triangleq \Sigma_{\boldsymbol{X}}^{-1}\Sigma_{\boldsymbol{XY}}\Sigma_{\boldsymbol{YX}}\Sigma_{\boldsymbol{X}}^{-1}$, which is always positive semidefinite. Since any positive semidefinite

$$\Phi \triangleq \mathbb{E}[(\boldsymbol{X}-\mathbb{E}[\boldsymbol{X}|\boldsymbol{Z}])(\boldsymbol{X}-\mathbb{E}[\boldsymbol{X}|\boldsymbol{Z}])^T]$$

is attainable via a linear encoding policy with a Gaussian perturbation, the problem reduces to

$$\begin{aligned} &\min_{\Phi\succeq 0}\ \operatorname{Tr}(\Upsilon\,\Phi) \\ &\text{subject to}\operatorname{Tr}(\Phi) \geq \alpha. \end{aligned} \tag{34}$$

Let the minimum eigenvalue of Υ be denoted by λ_{min}. Observe that

$$\begin{aligned} \operatorname{Tr}(\Upsilon\Phi) &= \operatorname{Tr}(\Upsilon\Phi - \lambda_{\text{min}}\Phi + \lambda_{\text{min}}\Phi) \\ &= \operatorname{Tr}((\Upsilon - \lambda_{\text{min}}I)\Phi) + \lambda_{\text{min}}\operatorname{Tr}(\Phi) \\ &\geq \alpha\lambda_{\text{min}} \end{aligned}$$

where the inequality uses the constraint along with the observation that $(\Upsilon - \lambda_{\text{min}}I)$ and Φ are positive semidefinite. As a result, by using a linear encoding policy possibly with a Gaussian perturbation, one can attain $\operatorname{Tr}(\Phi) = \alpha$ where the solution satisfies the orthogonality condition under the trace inner-product defining a Hilbert space on square matrices:

$$\operatorname{Tr}((\Upsilon - \lambda_{\text{min}}I)\Phi) = 0.$$

Since such an encoding policy achieves the characterized lower bound, it becomes the optimal solution to the constrained MMSE Gaussian information bottleneck problem. That the constrained problem with inequality is equivalent to a problem with an equality constraint applies more broadly to information bottleneck problems, see e.g. [58].

Remark 12: We emphasize that the solution presented in Theorem 3 is obtained without making an *a priori* linear policy restriction. These policies are the optimal solution among any set of policies for the optimization problem constructed at the sender by anticipating the best response of the receiver.

Remark 13: In the information bottleneck problem, the sender uses partial information since only random variable $\boldsymbol{X}$ is available at the sender whereas in our privacy-signaling game formulation the sender has access to both of the random variables. Due to this further restriction that only partial information is available at the sender, our information bottleneck analysis provides a lower bound on the performance of our original Stackelberg game setting.

Remark 14: It should be emphasized that the information bottleneck problem involving mutual information corresponds to the Stackelberg equilibrium concept since employing mutual information effectively means that the receiver uses all the available information, i.e., it employs its best response. On the other hand, the Nash problem would require an explicit dependence of the functions (considered in the optimization) on the receiver policy.

V. A Channel between the Sender and the Receiver

In this section, we generalize our results on the considered privacy-signaling game problem to scenarios when there is a channel between the sender and the receiver. In fact, the proposed equivalent formulation employed in the proof of Theorem 1 is also applicable when there is a channel between the sender and the receiver. Namely, Lemma 2 and Lemma 3 generalize to scenarios with a channel between the players represented by a conditional distribution $p(\boldsymbol{r}|\boldsymbol{z})$ where $\boldsymbol{R} = \boldsymbol{r}$ denotes the observation of the receiver. These generalizations imply that the sender cannot transmit information related to $\boldsymbol{V}$ and is restricted to send information related to $\boldsymbol{U}$ at a Nash equilibrium or at a Stackelberg equilibrium, where $\boldsymbol{U}$ and $\boldsymbol{V}$ are partitions of the transformed coordinate system specified in (14). Thus, for a given channel, the aim is to find an encoder/decoder pair that is optimal in conveying a sequence of independent zero-mean Gaussian distributed sources over that particular channel in a mean squared error sense and such an optimal encoding/decoding policy pair leads to a payoff dominant Nash equilibrium as well as a Stackelberg equilibrium.

In the following, we focus on the particular case of scalar sources and investigate the Nash and the Stackelberg equilibria for two important channel settings.

A. Gaussian Noise Channel between the Sender and the Receiver

In this subsection, we consider the same problem as before for scalar sources except that there is an additive Gaussian noise (e.g., measurement noise) between the transmitter and the receiver. More specifically, the sender encodes X and Y into Z which is subject to additive noise W and the receiver uses the observation $R = Z + W$ while decoding both of the random variables. The additive noise term is independent of X and Y and it is modeled as zero-mean Gaussian with variance σ_W^2. In addition, we assume that there is an average power constraint at the sender, i.e., $\mathbb{E}[Z^2] \leq P$.

1) Nash Equilibria:

Theorem 5:

(i) There exist informative linear Nash equilibria with an encoding policy $\gamma^e(x,y) = Ax + By$ that satisfies

$$\frac{B}{A} = -\frac{(\delta\sigma_X^2 + \sigma_Y^2) + \sqrt{(\delta\sigma_X^2 + \sigma_Y^2)^2 - 4\delta\rho^2}}{2\delta\rho}, \quad (35)$$

$$A^2\sigma_X^2 + B^2\sigma_Y^2 + 2AB\rho = P \quad (36)$$

and with decoding policies

$$\gamma^{d_X}(r) = \left(\frac{A\sigma_X^2 + B\rho}{P + \sigma_W^2}\right) r, \quad (37)$$

$$\gamma^{d_Y}(r) = \left(\frac{A\rho + B\sigma_Y^2}{P + \sigma_W^2}\right) r. \quad (38)$$

(ii) These informative equilibria are payoff dominant Nash equilibria and they are the only possible payoff dominant Nash equilibria. Moreover, these equilibria are unique among the affine class of policies.

Proof: Lemma 2 implies that the sender is restricted to convey U which corresponds to information conveyed by the sender at the equilibria specified in Corollary 1. Then, it is easy to verify that sending U after scaling up to the maximum available power level yields a Nash equilibrium. The decoding policies at this equilibrium are given by the minimum mean squared error estimators corresponding to each random variable. Since the observation R is jointly Gaussian with X and Y, the conditional expectation formula for Gaussian distributions can be employed to obtain (37) and (38) [62, p. 155].

The proof for the payoff dominance property of the equilibria uses the observation that the performance of both players is determined by $\mathbb{E}[(U - \gamma^{d_U}(R))^2]$ at a Nash equilibrium. Since the source is scalar and the Gaussian noise is additive, we can employ the well-known result that the problem of transmitting a scalar Gaussian source over a scalar Gaussian channel under an average power constraint admits a unique solution with linear encoding scaled to satisfy the power constraint with equality (see e.g. [64, p. 376]). Hence, the result immediately follows.

As it can be shown that an encoding policy $\tilde{\gamma}^e(u) = Au + C$ with $\mathbb{E}[(AU)^2] < P$ cannot be a Nash equilibrium, it follows that the only possible informative affine Nash equilibria are attained by an encoding policy that satisfies (35) and (36). ■

2) Stackelberg Equilibria:

Theorem 6: The Stackelberg equilibria coincide with the payoff dominant Nash equilibria characterized in Theorem 5. These equilibria are unique among any set of policies.

Proof: Lemma 3 implies that the encoder cannot convey V and it can only use U in constructing its message. As the objectives of each player then becomes the minimization of the mean squared error for estimating U, the optimal strategy of the sender is to employ an encoding policy which is linear in U with an average power equal to P. Moreover, this encoding strategy is unique due to the fact that it is the unique solution to the problem of transmitting a scalar Gaussian source over a scalar Gaussian channel under an average power constraint [64]. ■

It is important to emphasize that the encoder is not restricted to be affine. Since the problem reduces to transmitting a scalar Gaussian source over a scalar Gaussian channel under an average power constraint, we obtain these linear policies as the optimal unique solution to this reduced problem.

B. Discrete Noiseless Channel between the Sender and the Receiver

In this subsection, we consider scalar sources and investigate the discrete channel setting where the sender is restricted to transmit a discrete value, i.e., $Z \in \{0, \ldots, M-1\}$ for some $M \geq 2$. We assume that the channel is noiseless, i.e., $R = Z$.

While investigating the discrete channel setting, we again employ the equivalent formulation which facilitates the analysis. Lemma 2 and Lemma 3 imply that both players share the common objective of minimizing $\mathbb{E}[(U - \gamma^{d_U}(R))^2]$ under both of the equilibrium concepts. Since the sender is restricted to transmit discrete values, it is required to quantize

U at the sender. Since this would correspond to classical quantization, the existence of an optimal quantizer follows from the classical results in the literature, e.g., [65]. Namely, there exist quantization bins and reconstruction points which minimize the corresponding mean squared error. Thus, by assigning each bin to a discrete value of Z and then using the corresponding optimal reconstruction points at the receiver yield a Nash equilibrium. We summarize this result in the following theorem.

Theorem 7: Consider the quantization of U into M bins where each bin is assigned to a discrete value of Z at the encoder and the corresponding reconstruction points at the receiver such that $\mathbb{E}[(U - \gamma^{d_U}(R))^2]$ is minimized. This pair of encoding and decoding policies, which always exists, forms an informative Nash equilibrium. In addition, this equilibrium is a payoff dominant Nash equilibrium.

It is worth pointing out that for any number of bins lower than M, there exists a Nash equilibrium. In other words, even if Z can take M discrete values, a quantization policy using lower than M bins at the sender and the corresponding reconstruction points at the receiver is also a Nash equilibrium. In addition, the case of a single bin is also a Nash equilibrium where no information related to U is conveyed to the receiver.

It is noted that using a large number of bins yields a lower $\mathbb{E}[(U - \gamma^{d_U}(R))^2]$. Since the mean squared error for estimating U is desired to be minimized for both players, using a large number of bins results in improved objectives for both players. This monotonicity property with respect to the number of bins implies that at the Stackelberg equilibrium there must be M bins.

Theorem 8: The pair of policies in Theorem 7 leads to a Stackelberg equilibrium.

VI. Numerical Examples

In this section, we provide numerical examples for the proposed privacy-signaling game and the MMSE information bottleneck problems.

A. Scalar Sources

Here, we consider scalar sources and illustrate the performances at the characterized equilibria where the variances of the private and nonprivate random variables are set as $\sigma_X^2 = \sigma_Y^2 = 1$. We consider only the privacy-signaling game problem since the information bottleneck solution is simply given by the fully informative or noninformative solution depending on δ in the case of scalar sources as stated in Corollary 2. Since the informative Nash equilibrium coincides with the Stackelberg equilibrium in the case of scalar sources for the privacy-signaling game setup, we do not make a distinction between them.

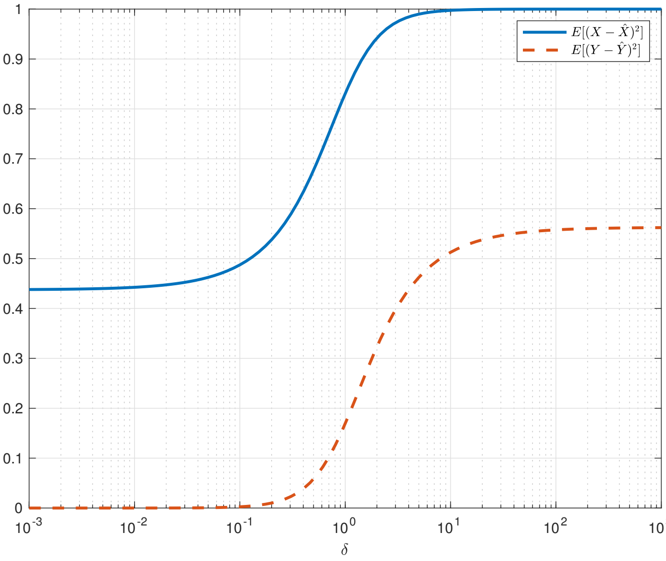


Fig. 3: Mean squared errors at the informative equilibrium for privacy-signaling game setup with respect to privacy ratio where $\rho = 0.75$.

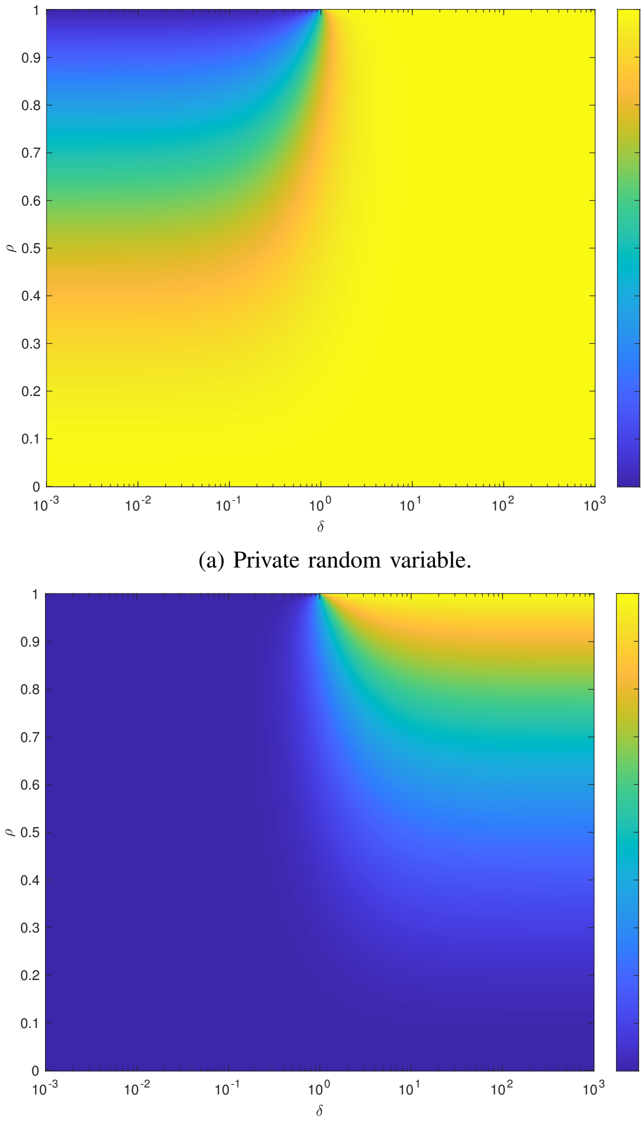
TABLE I: The ratio of coefficients at the encoder for the derived informative equilibria when $\sigma_X^2 = 1$ and $\sigma_Y^2 = 1$, as specified in (16).

Scenario	B/A
$\rho = 0.3$ and $\delta = 0.1$	-36.39
$\rho = 0.3$ and $\delta = 1$	-6.51
$\rho = 0.3$ and $\delta = 10$	-3.63
$\rho = 0.7$ and $\delta = 0.1$	-15.04
$\rho = 0.7$ and $\delta = 1$	-2.44
$\rho = 0.7$ and $\delta = 10$	-1.50

Fig. 3 plots the estimation errors for the private and nonprivate random variables with respect to the privacy ratio where the correlation between them is given by $\rho = 0.75$. The estimation error for the private random variable increases with the privacy ratio since the transmitter removes information related to the private random variable due to enhanced privacy concerns. This removal also distorts the information conveyed related to the nonprivate random variable and hence the corresponding estimation error also increases.

Next, we illustrate the attained costs with respect to both the privacy ratio δ and the correlation between the random variables ρ. We plot the estimation errors at the equilibria in Fig. 4a for the private random variable and in Fig. 4b for the nonprivate random variable. In the low privacy scenario, the estimation error for Y does not change significantly with respect to the correlation since most of the information contained in Y is conveyed to the receiver regardless of the correlation. As a result, more information is leaked related to the private random variable as the correlation is increased. In contrast, in the high privacy scenario, regardless of the correlation, most of the information related to X is removed from the transmitted message. Thus, the estimation error for the nonprivate random variable increases with the correlation whereas no significant changes in the estimation error for the private random variable are observed.

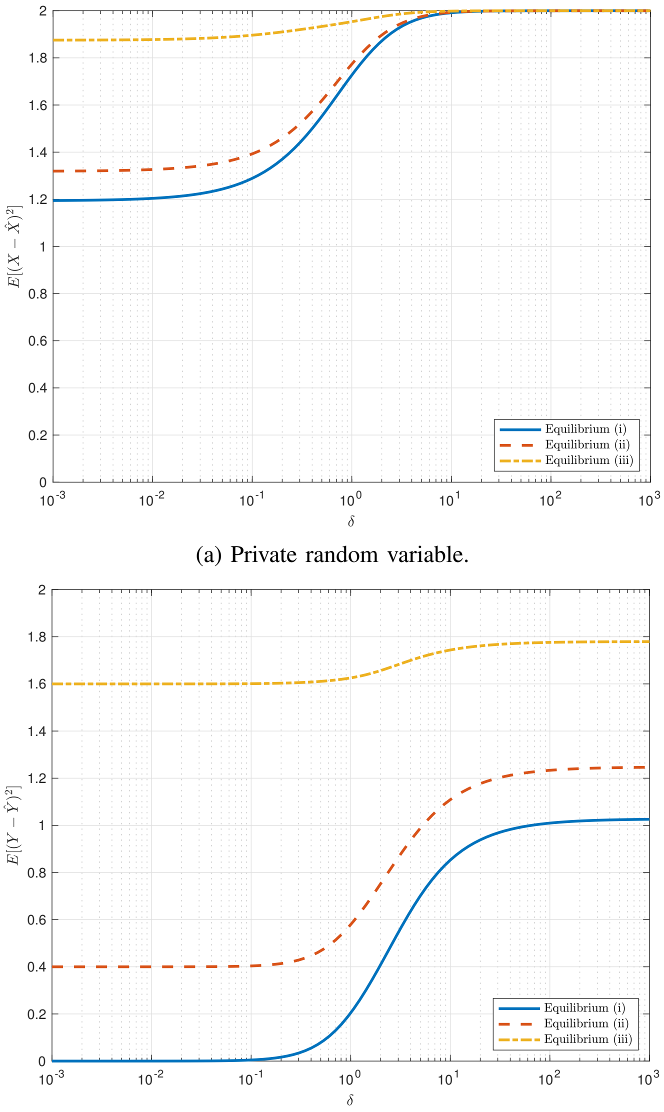
Table I illustrates the tradeoff between *utility* in terms of conveying Y and *privacy* in terms of hiding X by providing the structure of the encoder at the equilibrium for various values of the privacy ratio and correlation. It can be inferred that if the privacy ratio is increased while the correlation is kept the same, the information leakage related to the private



(a) Private random variable.

(b) Nonprivate random variable.

Fig. 4: Mean squared errors at the informative equilibrium for privacy-signaling game setup with respect to privacy ratio and correlation between the random variables.



(a) Private random variable.

(b) Nonprivate random variable.

Fig. 5: Mean squared errors at the informative equilibria for privacy-signaling game setup with respect to privacy ratio. The equilibrium (i) corresponds to the payoff dominant Nash equilibrium or the Stackelberg equilibrium, and the equilibria (ii) and (iii) correspond to two different Nash equilibria.

random variable reduces, as expected.

B. Multidimensional Sources

Here, we consider vector valued sources where both the private and the nonprivate random variables are two-dimensional with the following covariance matrix:

$$\Sigma = \begin{bmatrix} \Sigma_{\boldsymbol{X}} & \Sigma_{\boldsymbol{XY}} \\ \Sigma_{\boldsymbol{YX}} & \Sigma_{\boldsymbol{Y}} \end{bmatrix} = \begin{bmatrix} 1 & 0.7 & 0.7 & 0.6 \\ 0.7 & 1 & 0.2 & 0.5 \\ 0.7 & 0.2 & 1 & 0.6 \\ 0.6 & 0.5 & 0.6 & 1 \end{bmatrix}. \qquad (39)$$

We first illustrate the performance at the equilibria for the privacy-signaling game setup. Since both sources are multidimensional, there exist multiple linear Nash equilibria, which are characterized in Theorem 1. Among these Nash equilibria, one of them corresponds to the payoff dominant Nash equilibrium, which also coincides with the Stackelberg equilibrium as stated in Theorem 2. We plot the estimation errors at these informative equilibria with respect to the privacy ratio in Fig. 5a for the private random variable and in Fig. 5b for the nonprivate random variable. Similar to the scalar source setting, we observe that the estimation performance for both of the random variables degrades as δ increases since the sender removes more information related to the private random variable and thereby related to the nonprivate random variable. Moreover, the information conveyed at the payoff dominant Nash equilibria contains the information conveyed in other two Nash equilibria. Namely, the sender conveys both U_1 and U_2 at the payoff dominant Nash equilibria whereas the sender transmits U_1 or U_2 at the other two Nash equilibria

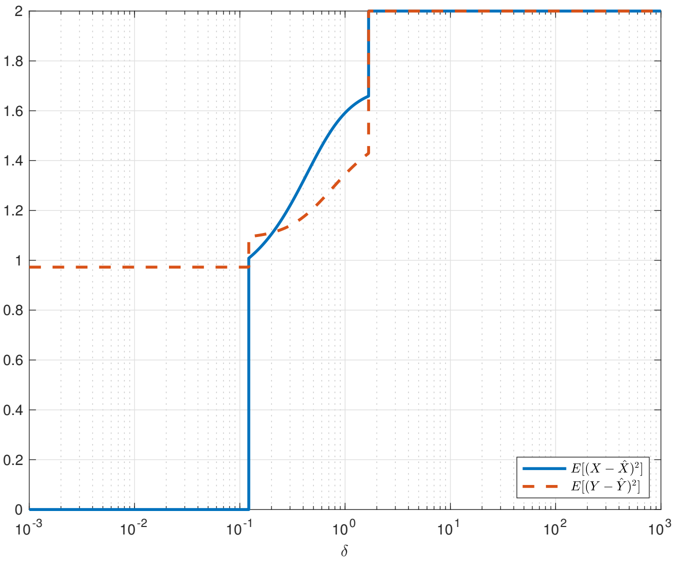


Fig. 6: Mean squared errors for the MMSE information bottleneck solution with respect to tradeoff parameter δ.

considering the transformed coordinate system defined by (10) and (14). It is seen that at the equilibrium (iii) in Fig. 5, the estimation errors do not change significantly with respect to δ in contrast to that at the equilibrium (ii). This reveals that for the considered setting the tradeoff between privacy and utility is more significant in one direction in the transformed coordinate system.

C. The MMSE Information Bottleneck Setup

Finally, we illustrate the performance at the MMSE information bottleneck solution. Fig. 6 plots the mean squared errors with respect to tradeoff parameter δ. The fully informative scenario with the sender revealing the random variable $\boldsymbol{X}$ completely is obtained for values of δ smaller than a certain threshold. In contrast, the solution becomes noninformative for values of δ larger than a certain threshold. When δ is between these values, the solution becomes informative with the sender applying a certain compression via a linear policy. It is interesting to observe jumps when δ is equal to the thresholds in Fig. 6. In fact, δ is equal to these thresholds when (31) has a term with a zero coefficient, i.e., $\lambda_j = 0$ for some $j \in \{1, \ldots, n_{\boldsymbol{X}}\}$. This implies that when δ is exactly equal to these thresholds, conveying the corresponding random variable T_j in the transformed coordinate system does not affect the sender's cost. As a result, we obtain the optimal solution when the sender transmits T_j as well as when the sender hides T_j completely or partially. Moreover, these thresholds for δ actually correspond to the values after which the dimension of the encoded message in equilibrium changes.

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 random variables have been modeled as jointly Gaussian random vectors. It has been proven that a payoff dominant Nash equilibrium is attained by linear policies. It has been shown that these linear policies at the payoff dominant Nash equilibria lead to Stackelberg equilibria as well. These results have been further generalized to the Gaussian noisy channel setting as well as a discrete noiseless channel setting for the special case of scalar sources. We have also provided an estimation theoretic perspective on the information bottleneck problem under the Stackelberg equilibrium concept. We have shown that the Stackelberg equilibria are attained by a set of characterized linear policies.

VIII. Acknowledgments

The authors would like to thank Prof. Tamer Başar for his detailed and insightful comments and pointing us to [27].

Appendix A
Supporting Results

In this appendix, we present supporting results used in the proofs of Theorem 1 and Theorem 2. In the proof of Theorem 1, we propose an equivalent formulation by introducing a linear transformation of variables. The following lemma establishes the optimality of the minimum mean squared error estimator at the decoder for a given encoding policy considering the proposed equivalent formulation.

Lemma 1: Consider the equivalent formulation illustrated in Fig. 2 where $\mathcal{T}$ and its inverse are fixed, and the encoder and the decoder select their corresponding policies $\tilde{\gamma}^e(\cdot)$ and $\gamma^{d_{\boldsymbol{T}}}(\cdot)$ arbitrarily. Then, for a fixed encoding function $\tilde{\gamma}^e(\boldsymbol{t})$, the optimal $\gamma^{d_{\boldsymbol{T}}}(\boldsymbol{z})$ that minimizes (13) is given by $\mathbb{E}[\boldsymbol{T}|\boldsymbol{Z} = \boldsymbol{z}]$.

Proof: The result is standard but for completeness we present a short proof. Suppose that $\gamma^{d_{\boldsymbol{T}}}(\boldsymbol{z}) = \mathbb{E}[\boldsymbol{T}|\boldsymbol{Z} = \boldsymbol{z}] + g(\boldsymbol{z})$. Inserting this expressions into the objective function of the receiver in (13), we get

$$
\begin{aligned}
J^d(g) = \mathbb{E}[&(\boldsymbol{T} - \mathbb{E}[\boldsymbol{T}|\boldsymbol{Z}] - g(\boldsymbol{Z}))^T K \\
&(\boldsymbol{T} - \mathbb{E}[\boldsymbol{T}|\boldsymbol{Z}] - g(\boldsymbol{Z}))] \\
= \mathbb{E}&[(\boldsymbol{T} - \mathbb{E}[\boldsymbol{T}|\boldsymbol{Z}])^T K (\boldsymbol{T} - \mathbb{E}[\boldsymbol{T}|\boldsymbol{Z}])] \\
&+ \mathbb{E}[g(\boldsymbol{Z})^T K g(\boldsymbol{Z})] \\
\geq \mathbb{E}&[(\boldsymbol{T} - \mathbb{E}[\boldsymbol{T}|\boldsymbol{Z}])^T K (\boldsymbol{T} - \mathbb{E}[\boldsymbol{T}|\boldsymbol{Z}])]
\end{aligned}
$$

where the inequality follows from $K = Q^T \Sigma Q$ being positive definite. This proves the optimality of the minimum mean squared error estimator in the transformed coordinate system for a given encoding policy. ■

In the following lemma, we show that the sender can only transmit information related to one of the random variables at a Nash equilibrium considering the proposed equivalent formulation illustrated in Fig.2.

Lemma 2: Consider the privacy-signaling game problem. At a Nash equilibrium, the sender does not reveal any information related to the linear combinations (of the private and nonprivate random variables) $\boldsymbol{V}$ specified in (14).

Proof: Consider a set of policies where the sender employs an encoding policy $\tilde{\gamma}^e(\boldsymbol{t}) = f(\boldsymbol{t})$ which conveys information related to T_j for some j with $\lambda_j < 0$. In response to this encoding policy, it is optimal for the receiver to employ the minimum mean squared error estimators of each random variable, as shown in Lemma 1. Denote these estimators for estimating T_i by $g_i(\boldsymbol{z})$ for $i = 1, \dots, n$. Since we assume that the encoding policy $f(\boldsymbol{t})$ conveys information related to T_j, the mean squared error for estimating T_j with the corresponding optimal estimator $g_j(\boldsymbol{z})$ is lower than $\sigma^2_{T_j}$, i.e., $\mathbb{E}[(T_j - g_j(\boldsymbol{Z}))^2] < \sigma^2_{T_j}$. In response to the decoding policies of $\{g_i(\cdot)\}_{i=1}^n$, the sender can switch to the following policy to improve its objective value. Instead of sending $\boldsymbol{z} = f(\boldsymbol{t})$, the sender can transmit $\boldsymbol{z} = f(\bar{\boldsymbol{t}})$ while keeping the encoding function $f(\cdot)$ the same where $\bar{\boldsymbol{t}} \triangleq [t_1, \dots, t_{j-1}, w, t_{j+1}, \dots, t_n]^T$ and w is a realization of a random variable that follows the same distribution as T_j and is independent of $\boldsymbol{T}$. In that case, the performance for estimating T_i for $i \neq j$ remains the same since receiving $f(\bar{\boldsymbol{t}})$ or $f(\boldsymbol{t})$ are equivalent for the decoding policy $g_i(\boldsymbol{z})$. However, the performance for estimating T_j degrades as shown in the following:

$$\begin{aligned}\mathbb{E}[(T_j - g_j(\boldsymbol{Z}))^2] &= \mathbb{E}[T_j^2] - 2\mathbb{E}[T_j g_j(\boldsymbol{Z})] + \mathbb{E}[g_j(\boldsymbol{Z})^2] \\ &= \mathbb{E}[T_j^2] - 2\mathbb{E}[T_j]\mathbb{E}[g_j(\boldsymbol{Z})] + \mathbb{E}[g_j(\boldsymbol{Z})^2] \\ &= \mathbb{E}[T_j^2] + \mathbb{E}[g_j(\boldsymbol{Z})^2] \geq \sigma^2_{T_j},\end{aligned}$$

where the second equality is due to the fact that T_j and $\boldsymbol{Z}$ are independent in case $f(\bar{\boldsymbol{T}})$ is transmitted. Since the random variable T_j is chosen such that $\lambda_j < 0$ in (12), the sender gains by employing $\boldsymbol{z} = f(\bar{\boldsymbol{t}})$ instead of $\boldsymbol{z} = f(\boldsymbol{t})$. As a result, any encoding policy which yields $\mathbb{E}[(T_j - \gamma^{d_{T_j}}(\boldsymbol{Z}))^2] < \sigma^2_{T_j}$ for an index j with $\lambda_j < 0$ cannot be a Nash equilibrium since in that case the sender can change its strategy to improve its objective value.

In game theory, when a unilateral change by a decision maker occurs, the perturbed policies may cease to be an equilibrium. However, a subtle aspect of our proof is that, the revised sender policy does not alter the policy of the decoder, therefore the perturbation is still an equilibrium. ■

Similar to the result of Lemma 2 which applies to a Nash equilibrium, the sender is restricted to transmit information related to $\boldsymbol{U}$ at a Stackelberg equilibrium. The following lemma proves this result.

Lemma 3: Consider the privacy-signaling game problem. At a Stackelberg equilibrium, the sender does not reveal any information related to the linear combinations (of the private and nonprivate random variables) $\boldsymbol{V}$ specified in (14).

Proof: We show that any encoding policy which yields $\mathbb{E}[(T_j - \gamma^{d_{T_j}}(\boldsymbol{Z}))^2] < \sigma^2_{T_j}$ for an index j with $\lambda_j < 0$ cannot be a Stackelberg equilibrium via a similar analysis to that employed in Lemma 2. Towards that goal, we compare the performance of two scenarios from the perspective of the sender. Recall that in a Stackelberg equilibrium the sender chooses a policy and announces this policy to the receiver and the receiver acts with the knowledge of sender's policy. Denote the encoding policy by $\tilde{\gamma}^e(\boldsymbol{t}) = f(\boldsymbol{t})$ in the first scenario. The receiver takes an optimal response to this announced encoding policy. Assume that $\mathbb{E}[(T_j - \gamma^{d_{T_j}}(\boldsymbol{Z}))^2] < \sigma^2_{T_j}$ with the corresponding set of policies. In the second scenario, suppose that the encoder chooses the same policy as before with the exception that the sender replaces the realization $T_j = t_j$ by an independent noise following the same distribution as T_j. Namely, the sender uses $f(\bar{\boldsymbol{t}})$ where $\bar{\boldsymbol{t}} = [t_1, \dots, t_{j-1}, w, t_{j+1}, \dots, t_n]^T$ and w is a realization of a random variable that follows the same distribution as T_j and is independent of $\boldsymbol{T}$. As the sender announces its strategy, the optimal response of the receiver for the random variable T_j becomes $\gamma^{d_{T_j}}(\boldsymbol{z}) = \mathbb{E}[T_j|\boldsymbol{Z} = \boldsymbol{z}] = \mathbb{E}[T_j] = 0$ due to the independence of T_j and $\boldsymbol{Z}$ in this scenario. Therefore, we get $\mathbb{E}[(T_j - \gamma^{d_{T_j}}(\boldsymbol{Z}))^2] = \sigma^2_{T_j}$ in this case. Notice that the mean squared error performance in estimating T_i for $i \neq j$ is the same for both scenarios. As a result, the second scenario yields better performance for the sender. Thus, transmitting information related to T_j with $\lambda_j < 0$ in (12) is not desirable for the sender. ■

References

[1] E. Kazıklı, S. Gezici, and S. Yüksel, "Quadratic privacy-signaling games and payoff dominant equilibria," in *IEEE International Symposium on Information Theory (ISIT)*, 2020, pp. 1367–1372.

[2] P. McDaniel and S. McLaughlin, "Security and privacy challenges in the smart grid," *IEEE Security Privacy*, vol. 7, no. 3, pp. 75–77, May 2009.

[3] S. Finster and I. Baumgart, "Privacy-aware smart metering: A survey," *IEEE Communications Surveys & Tutorials*, vol. 17, no. 2, pp. 1088–1101, 2015.

[4] S. Han, U. Topcu, and G. J. Pappas, "Event-based information-theoretic privacy: A case study of smart meters," in *American Control Conference (ACC)*, 2016, pp. 2074–2079.

[5] 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 (SmartGridComm)*, 2011, pp. 220–225.

[6] J. Yao and P. Venkitasubramaniam, "On the privacy-cost tradeoff of an in-home power storage mechanism," in *Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, 2013, pp. 115–122.

[7] D. He, S. Chan, and M. Guizani, "User privacy and data trustworthiness in mobile crowd sensing," *IEEE Wireless Communications*, vol. 22, no. 1, pp. 28–34, Feb. 2015.

[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, no. 5, pp. 839–853, Oct. 2016.

[9] 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, no. 6, pp. 918–923, Nov. 1983.

[10] S. Asoodeh, M. Diaz, F. Alajaji, and T. Linder, "Estimation efficiency under privacy constraints," *IEEE Transactions on Information Theory*, vol. 65, no. 3, pp. 1512–1534, March 2019.

[11] H. Wang, L. Vo, F. P. Calmon, M. Médard, K. R. Duffy, and M. Varia, "Privacy with estimation guarantees," *IEEE Transactions on Information Theory*, vol. 65, no. 12, pp. 8025–8042, Dec. 2019.

[12] A. Padakandla, P. R. Kumar, and W. Szpankowski, "The trade-off between privacy and fidelity via Ehrhart theory," *IEEE Transactions on Information Theory*, vol. 66, no. 4, pp. 2549–2569, Apr. 2020.

[13] M. Diaz, H. Wang, F. P. Calmon, and L. Sankar, "On the robustness of information-theoretic privacy measures and mechanisms," *IEEE Transactions on Information Theory*, vol. 66, no. 4, pp. 1949–1978, Apr. 2020.

[14] 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, no. 6, pp. 838–852, June 2013.

[15] 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, March 2020.

[16] S. Sreekumar and D. Gündüz, "Optimal privacy-utility trade-off under a rate constraint," in *IEEE International Symposium on Information Theory (ISIT)*, 2019, pp. 2159–2163.

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

[18] Y. Lu and M. Zhu, “On privacy preserving data release of linear dynamic networks,” *Automatica*, vol. 115, p. 108839, May 2020.

[19] J. Le Ny and G. J. Pappas, “Differentially private filtering,” *IEEE Transactions on Automatic Control*, vol. 59, no. 2, pp. 341–354, Feb. 2014.

[20] Y. Mo and R. M. Murray, “Privacy preserving average consensus,” *IEEE Transactions on Automatic Control*, vol. 62, no. 2, pp. 753–765, Feb. 2017.

[21] E. Nekouei, T. Tanaka, M. Skoglund, and K. H. Johansson, “Information-theoretic approaches to privacy in estimation and control,” *Annual Reviews in Control*, vol. 47, pp. 412–422, 2019.

[22] Y. Wang, Z. Huang, S. Mitra, and G. E. Dullerud, “Differential privacy in linear distributed control systems: Entropy minimizing mechanisms and performance tradeoffs,” *IEEE Transactions on Control of Network Systems*, vol. 4, no. 1, pp. 118–130, March 2017.

[23] T. Başar and G. J. Olsder, *Dynamic Noncooperative Game Theory*. Philadelphia, PA: SIAM Classics in Applied Mathematics, 1999.

[24] E. Kamenica and M. Gentzkow, “Bayesian persuasion,” *American Economic Review*, vol. 101, no. 6, pp. 2590–2615, Oct. 2011.

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

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

[27] W. Tamura, “Bayesian persuasion with quadratic preferences,” 2018, working paper, available at SSRN 1987877.

[28] 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, Feb. 2017.

[29] S. Sarıtaş, 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, Feb. 2017.

[30] 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, Feb. 2017.

[31] V. S. S. Nadendla, C. Langbort, and T. Başar, “Effects of subjective biases on strategic information transmission,” *IEEE Transactions on Communications*, vol. 66, no. 12, pp. 6040–6049, Dec. 2018.

[32] M. Le Treust and T. Tomala, “Persuasion with limited communication capacity,” *Journal of Economic Theory*, vol. 184, p. 104940, Nov. 2019.

[33] ——, “Information-theoretic limits of strategic communication,” *arXiv: 1807.05147*, 2018.

[34] S. Sarıtaş, S. Yüksel, and S. Gezici, “Dynamic signaling games with quadratic criteria under Nash and Stackelberg equilibria,” *Automatica*, vol. 115, p. 108883, May 2020.

[35] E. Kazıklı, S. Sarıtaş, S. Gezici, T. Linder, and S. Yüksel, “Signaling games for log-concave distributions: Number of bins and properties of equilibria,” *IEEE Transactions on Information Theory*, vol. 68, no. 3, pp. 1731–1757, March 2022.

[36] M. O. Sayın, E. Akyol, and T. Başar, “Hierarchical multistage Gaussian signaling games in noncooperative communication and control systems,” *Automatica*, vol. 107, pp. 9–20, Sep. 2019.

[37] M. O. Sayın and T. Başar, “Bayesian persuasion with state-dependent quadratic cost measures,” *IEEE Transactions on Automatic Control*, vol. 67, no. 3, pp. 1241–1252, March 2022.

[38] E. Kazıklı, S. Gezici, and S. Yüksel, “Signaling games in higher dimensions: Geometric properties of equilibrium solutions,” *arXiv: 2108.05240*, 2021.

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

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

[41] F. Farokhi and G. Nair, “Privacy-constrained communication,” *IFAC-PapersOnLine*, vol. 49, no. 22, pp. 43–48, Nov. 2016.

[42] E. Akyol, C. Langbort, and T. Başar, “Strategic compression and transmission of information,” in *IEEE Information Theory Workshop - Fall (ITW)*, 2015, pp. 219–223.

[43] 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. Springer, 2008, pp. 1–19.

[44] 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.

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

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

[47] S. Han, U. Topcu, and G. J. Pappas, “Differentially private distributed constrained optimization,” *IEEE Transactions on Automatic Control*, vol. 62, no. 1, pp. 50–64, Jan. 2017.

[48] J. Le Ny, *Differential Privacy for Dynamic Data*. Springer, 2020.

[49] N. Tishby and N. Zaslavsky, “Deep learning and the information bottleneck principle,” in *IEEE Information Theory Workshop (ITW)*, 2015, pp. 1–5.

[50] P. Harremoës and N. Tishby, “The information bottleneck revisited or how to choose a good distortion measure,” in *IEEE International Symposium on Information Theory*, 2007, pp. 566–570.

[51] R. Gilad-Bachrach, A. Navot, and N. Tishby, “An information theoretic tradeoff between complexity and accuracy,” in *Learning Theory and Kernel Machines*, B. Schölkopf and M. K. Warmuth, Eds. Springer, 2003, pp. 595–609.

[52] I. S. Dhillon, S. Mallela, and D. S. Modha, “Information-theoretic co-clustering,” in *Proceedings of the Ninth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2003, p. 89–98.

[53] M. Vera, P. Piantanida, and L. R. Vega, “The role of the information bottleneck in representation learning,” in *IEEE International Symposium on Information Theory (ISIT)*, 2018, pp. 1580–1584.

[54] M. Vera, L. R. Vega, and P. Piantanida, “Collaborative information bottleneck,” *IEEE Transactions on Information Theory*, vol. 65, no. 2, pp. 787–815, Feb. 2019.

[55] Z. Goldfeld and Y. Polyanskiy, “The information bottleneck problem and its applications in machine learning,” *IEEE Journal on Selected Areas in Information Theory*, vol. 1, no. 1, pp. 19–38, May 2020.

[56] H. Hsu, S. Asoodeh, S. Salamatian, and F. P. Calmon, “Generalizing bottleneck problems,” in *IEEE International Symposium on Information Theory (ISIT)*, 2018, pp. 531–535.

[57] A. Zaidi, I. Estella-Aguerri, and S. Shamai (Shitz), “On the information bottleneck problems: Models, connections, applications and information theoretic views,” *Entropy*, vol. 22, no. 2, Jan. 2020.

[58] H. Witsenhausen and A. Wyner, “A conditional entropy bound for a pair of discrete random variables,” *IEEE Transactions on Information Theory*, vol. 21, no. 5, pp. 493–501, Sep. 1975.

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

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

[61] R. A. Horn and C. R. Johnson, *Matrix Analysis*. New York: Cambridge University Press, 2013.

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

[63] A. Globerson and N. Tishby, “On the optimality of Gaussian information bottleneck curve,” Hebrew Univ, Jerusalem, Israel, Tech. Rep., 2004.

[64] S. Yüksel and T. Başar, *Stochastic Networked Control Systems: Stabilization and Optimization under Information Constraints*. Boston, MA: Birkhauser, 2013.

[65] R. M. Gray and D. L. Neuhoff, “Quantization,” *IEEE Transactions on Information Theory*, vol. 44, no. 6, pp. 2325–2383, Oct. 1998.

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 the École Polytechnique Fédérale de Lausanne (EPFL), Switzerland, in 2015, and the Ph.D. degree in electrical and electronics engineering from Bilkent University in 2021. He is currently a Post-Doctoral Researcher at the Department of Mathematics and Statistics, Queen’s University, Canada. His research interests include game theory, communication theory, and information theory.

Sinan Gezici (Senior Member, IEEE) 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, 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 (Member, IEEE) 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.