

Dynamic Signaling Games under Nash and Stackelberg Equilibria

Serkan Sarıtaş

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

Serdar Yüksel

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

Sinan Gezici

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

Abstract—In this study, dynamic and repeated quadratic cheap talk and signaling game problems are investigated. These involve encoder and decoders with mismatched performance objectives, where the encoder has a bias term in the quadratic cost functional. We consider both Nash equilibria and Stackelberg equilibria as our solution concepts, under a perfect Bayesian formulation. These two lead to drastically different characteristics for the equilibria. For the cheap talk problem under Nash equilibria, we show that fully revealing equilibria cannot exist and the final state equilibria have to be quantized for a large class of source models; whereas, for the Stackelberg case, the equilibria must be fully revealing regardless of the source model. In the dynamic signaling game where the transmission of a Gaussian source over a Gaussian channel is considered, the equilibrium policies are always linear for scalar sources under Stackelberg equilibria, and affine policies constitute an invariant subspace under best response maps for Nash equilibria.

I. INTRODUCTION

Signaling games and cheap talk are concerned with a class of Bayesian games where an informed player (encoder, sender) transmits information to another player (decoder, receiver). What makes such problems different from the classical communication problems is that the objective functions of the encoder and the decoder are not identical; in the cheap talk setup the cost functions do not depend on the transmitted signals and in the signaling game the cost functions may explicitly depend on the transmitted signals. The cheap talk problem was studied by Crawford and Sobel [1], who obtained the striking result that under some technical conditions on the cost functions, the cheap talk problem only admits equilibria that are essentially quantization policies. This is in contrast with the case where the goals are aligned [2].

The cheap talk and signaling game problems find applications in networked control systems when a communication channel/network is present among competitive and non-cooperative decision makers [3], [4]. Also, there have been a number of related contributions in the economics literature in addition to the seminal work by Crawford and Sobel, which are reviewed in [5]. [6] considers a Gaussian cheap talk game with quadratic cost functions where the analysis considers perfect Bayesian (Stackelberg) equilibria, for a class of single- and multi-terminal setups and

where affine equilibria are studied. In [7], it is shown that the Stackelberg equilibrium strategies are affine in the quadratic Gaussian cheap talk.

In our previous work [5], we consider both (simultaneous) Nash equilibria and (sequential) Stackelberg equilibria for the setup of Crawford and Sobel under quadratic cost criteria where the encoder has an additive bias term in his cost function, and its multi-dimensional and noisy extensions. We show that for arbitrary scalar sources, the quantized nature of all equilibrium policies holds under Nash equilibria, whereas all Stackelberg equilibria policies are fully informative. For multi-dimensional setups, unlike the scalar case, Nash equilibrium policies may be of non-quantized nature, and even linear. In the noisy setup, we present conditions for the existence of affine Nash equilibria as well as general informative equilibria and show that the only Stackelberg equilibrium is the linear equilibrium when the variables are scalar.

The dynamic (multi-stage) extension of the setup of Crawford and Sobel has been analyzed under various setups in the economics literature. In [8] and [9], the dynamic cheap talk is studied where the former one assumes that the source is a fixed random variable distributed according to some density on $[0, 1]$, and the latter one assumes that the sequence of states follows an irreducible Markov chain, and the set of states, messages and actions are finite.

A. Contributions in this Study

This study focuses on the multi-stage setup of a cheap talk problem introduced by Crawford and Sobel [1]. We extend the static cheap talk and signaling game studied in our previous work [5] to a dynamic setup. We prove that, in a repeated cheap talk game, under Nash equilibria, the last stage equilibria are quantized for any source with arbitrary distribution and fully revealing equilibria cannot exist for a class of models (see Remark 3.1), whereas the equilibrium must be fully revealing in the dynamic Stackelberg cheap talk game. The dynamic signaling game is a noisy version of the dynamic cheap talk game; i.e., for each stage, a scalar Gaussian source is to be transmitted over an additive Gaussian channel and the goals of the encoder and the decoder are misaligned by a bias term and encoder's cost also includes a penalty term for the transmitted signal. Under Nash equilibria, it is shown that the encoder (decoder) must be affine for an affine decoder

This research was supported in part by the Natural Sciences and Engineering Research Council (NSERC) of Canada, The Scientific and Technological Research Council of Turkey (TÜBİTAK) and the Distinguished Young Scientist Award of Turkish Academy of Sciences (TÜBA-GEBİP 2013).

(encoder); whereas, the only equilibrium in the dynamic Stackelberg signaling game is the linear equilibrium.

II. PRELIMINARIES

Let there be an encoder who wishes to encode the \mathbb{M} -valued random variable M and transmits \mathbb{X} -valued random variable X to a decoder. The decoder, upon receiving X , generates its optimal decision U which is also taken to be \mathbb{M} -valued. We consider here only deterministic encoder and decoder policies; i.e., $x = \gamma^e(m)$ and $u = \gamma^d(x) = \gamma^d(\gamma^e(m))$.

The encoder aims to find an optimal γ^e that minimizes

$$J^e(\gamma^e, \gamma^d) = \int c^e(m, u)P(dm),$$

whereas the decoder aims to find an optimal γ^d that minimizes

$$J^d(\gamma^e, \gamma^d) = \int c^d(m, u)P(dm).$$

Such a problem is known in the economics literature as *cheap talk* (since the transmitted signal does not affect the cost). A more general formulation would be the case when the transmitted signal is also an explicit part of the cost function c^e or c^d ; in that case, the setup is called a *signaling game*.

Since the goals are not aligned, such a problem is studied under the tools and concepts provided by *game theory*. A pair of policies $\gamma^{*,e}, \gamma^{*,d}$ is said to be a (simultaneous) **Nash equilibrium** if

$$\begin{aligned} J^e(\gamma^{*,e}, \gamma^{*,d}) &\leq J^e(\gamma^e, \gamma^{*,d}) \quad \forall \gamma^e \in \Gamma^e \\ J^d(\gamma^{*,e}, \gamma^{*,d}) &\leq J^d(\gamma^{*,e}, \gamma^d) \quad \forall \gamma^d \in \Gamma^d \end{aligned} \quad (1)$$

where Γ^e and Γ^d are the set of all deterministic functions from \mathbb{M} to \mathbb{X} and from \mathbb{X} to \mathbb{M} , respectively.

In the discussion so far, a simultaneous game-play is assumed and thus equilibrium refers to a Nash equilibrium. Besides the simultaneous game-play, one can also consider a sequential game-play; i.e. first the encoder sends the message and announces his/her policy, then the decoder receives them and takes an action sequentially, which leads to **Stackelberg** equilibria. In the Stackelberg game, the encoder announces his/her coding strategy and since the decoder takes an action after receiving the message, the encoder knows the optimal action which will be taken by the decoder and chooses the message to be transmitted accordingly. A pair of policies $\gamma^{*,e}, \gamma^{*,d}$ is said to be a **Stackelberg equilibrium** if

$$\begin{aligned} J^e(\gamma^{*,e}, \gamma^{*,d}(\gamma^{*,e})) &\leq J^e(\gamma^e, \gamma^{*,d}(\gamma^e)) \quad \forall \gamma^e \in \Gamma^e \\ \text{where } \gamma^{*,d}(\gamma^e) &\text{ satisfies} \\ J^d(\gamma^e, \gamma^{*,d}(\gamma^e)) &\leq J^d(\gamma^e, \gamma^d(\gamma^e)) \quad \forall \gamma^d \in \Gamma^d. \end{aligned} \quad (2)$$

All the game setups described above are *static (one-stage) game*. If a game is played over a number of time periods, the game is called a *dynamic game*. In each stage of the game, say stage- k , the encoder wishes to encode the \mathbb{M} -valued random variable M_k to the decoder by knowing the values of $\mathcal{I}_k^e = \{m_{[0,k]}, x_{[0,k-1]}\}$ with $\mathcal{I}_0^e = \{m_0\}$ where $m_{[0,k]} = \{m_0, m_1, \dots, m_k\}$ and $x_{[0,k-1]} = \{x_0, x_1, \dots, x_{k-1}\}$. Let X_k denote the \mathbb{X} -valued random variable which is transmitted to the

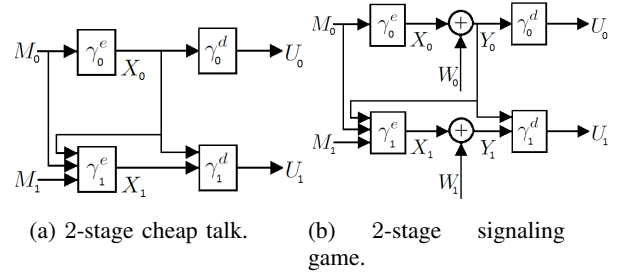


Fig. 1: General system model

decoder. The decoder, upon receiving X_k , generates its optimal decision U_k which is also \mathbb{M} -valued by knowing the values of $\mathcal{I}_k^d = \{x_{[0,k]}\}$. Thus, under the policies considered, $x_k = \gamma_k^e(\mathcal{I}_k^e)$ and $u_k = \gamma_k^d(\mathcal{I}_k^d)$.

The goal of the encoder is to find a policy sequence $\gamma^e = \{\gamma_0^e, \gamma_1^e, \dots, \gamma_{N-1}^e\}$ that minimizes

$$J^e(\gamma^e, \gamma^d) = \sum_{k=0}^{N-1} \int c_k^e(m_k, u_k)P(dm_k), \quad (3)$$

whereas the goal of the decoder is to find a policy sequence $\gamma^d = \{\gamma_0^d, \gamma_1^d, \dots, \gamma_{N-1}^d\}$ that minimizes

$$J^d(\gamma^e, \gamma^d) = \sum_{k=0}^{N-1} \int c_k^d(m_k, u_k)P(dm_k). \quad (4)$$

Using the encoder cost in (3) and the decoder cost in (4), the Nash equilibrium and the Stackelberg equilibrium for dynamic games can be defined in the same way as in (1) and (2), respectively.

Under both equilibria concepts, we consider the setups where the decision makers act optimally for each history path of the game (available to each decision maker) and the updates are Bayesian, and thus the equilibria form *Perfect Bayesian Equilibria*.

In this study, the quadratic cost functions are assumed; i.e., $c_k^e(m_k, u_k) = (m_k - u_k - b)^2$ and $c_k^d(m_k, u_k) = (m_k - u_k)^2$ where b is the bias term as in [1] and [5].

III. DYNAMIC CHEAP TALK

For the purpose of illustration, the system model of the 2-stage dynamic cheap talk is shown in Fig. 1-a.

A. Repeated Static Games: Nash Equilibria

In this part, the dynamic cheap talk game with an i.i.d. source will be analyzed. Since the source is assumed to be i.i.d., the game is a *repeated game*. First, the results on the 2-stage repeated games will be presented, then these results will be extended to the N -stage repeated games.

Theorem 3.1: Assuming the deterministic equilibrium for the 2-stage repeated cheap talk game, the equilibrium policies for the second stage must be quantized almost surely for any collection of policies (γ_0^e, γ_0^d) and for any real-valued source model with arbitrary probability measure.

Proof: Given the second stage encoder and decoder policies $\gamma_1^e(m_0, m_1, x_0)$ and $\gamma_1^d(x_0, x_1)$, it is possible to define policies which are parametrized by the common information x_0 almost surely so that $\hat{\gamma}_{x_0}^e(m_0, m_1) = \gamma_1^e(m_0, m_1, x_0)$ and $\hat{\gamma}_{x_0}^d(x_1) = \gamma_1^d(x_0, x_1)$.

Now fix the first stage policies γ_0^e and γ_0^d . Suppose that the second stage encoder does not use m_0 ; i.e., $\hat{\gamma}_{x_0}^{e1}(m_1)$ is the policy of the second stage encoder. For the policies $\hat{\gamma}_{x_0}^{e1}(m_1)$ and $\hat{\gamma}_{x_0}^d(x_1)$, by using the second stage encoder cost function $F_{x_0}(m_1, u_1) \triangleq E[(m_1 - u_1 - b)^2 | x_0]$ and the bin arguments from [5, Theorem 3.2], it can be deduced that the equilibrium policies for the second stage must be quantized, for any collection of policies (γ_0^e, γ_0^d) and for any given x_0 due to the continuity of $F_{x_0}(m_1, u_1)$ in m_1 . Now let the second stage encoder use m_0 ; i.e., $\hat{\gamma}_{x_0}^e(m_0, m_1)$ is the policy of the second stage encoder. Here, even if $\hat{\gamma}_{x_0}^e(m_0, m_1)$ is a deterministic policy, it can be seen as an equivalent randomized encoder policy (as a stochastic kernel from \mathbb{M}_1 to \mathbb{X}_1) where m_0 is a real valued random variable independent of m_1 . In [5], we considered randomized encoders as well and thus from [5, Theorem 3.3], the equilibrium is achievable with an encoder policy which uses only m_1 ; i.e., $\hat{\gamma}_{x_0}^{e*}(m_1)$ is an encoder policy at the equilibrium and thus the equilibria are quantized. ■

Theorem 3.2: For the 2-stage repeated cheap talk game, if the sources m_0 and m_1 are uniform on $[0, 1]$, then the first stage equilibrium cannot be a fully revealing equilibrium.

Proof: Let two bins of the first stage equilibrium be \mathcal{B}_0^α and \mathcal{B}_0^β , and their encoding values be x_0^α and x_0^β , respectively. Also let m_0^α indicate any point in \mathcal{B}_0^α ; i.e., $m_0^\alpha \in \mathcal{B}_0^\alpha$. Similarly, let m_0^β represent any point in \mathcal{B}_0^β ; i.e., $m_0^\beta \in \mathcal{B}_0^\beta$. The decoder chooses action $u_0^\alpha = \gamma_0^d(x_0^\alpha)$ when the encoder sends $x_0^\alpha = \gamma_0^e(m_0^\alpha)$ and action $u_0^\beta = \gamma_0^d(x_0^\beta)$ when the encoder sends $x_0^\beta = \gamma_0^e(m_0^\beta)$ in order to minimize its total cost.

Let $F(m_0, x_0)$ be a cost function for the first stage encoder if it encodes message m_0 as x_0 . Since the second stage equilibrium cost does not depend on m_0 by Theorem 3.1, $F(m_0, x_0)$ can be written as $F(m_0, x_0) = (m_0 - \gamma_0^d(x_0) - b)^2 + G(x_0)$ where $G(x_0) = E_{m_1} \left[\left(m_1 - \gamma_1^{*,d}(x_0, \gamma_1^{*,e}(m_1, x_0)) - b \right)^2 \middle| x_0 \right]$ is the expected cost of the second stage encoder, and $\gamma_1^{*,e}$ and $\gamma_1^{*,d}$ are the second stage encoder and decoder policies at the equilibrium, respectively. Since under any equilibrium, the maximum number of bins is finite when the source is uniform on $[0, 1]$, there are finitely many equilibria at the second stage which implies that the second stage encoder cost can take finitely many different values; i.e., $G(x_0)$ can take finitely many values.

Due to the equilibrium definitions from the view of the encoder, $F(m_0^\alpha, x_0^\alpha) < F(m_0^\alpha, x_0^\beta)$ and $F(m_0^\beta, x_0^\beta) < F(m_0^\beta, x_0^\alpha)$. These inequalities imply that

$$\begin{aligned} (m_0^\alpha - u_0^\alpha - b)^2 + G(x_0^\alpha) &< (m_0^\alpha - u_0^\beta - b)^2 + G(x_0^\beta) \\ (m_0^\beta - u_0^\beta - b)^2 + G(x_0^\beta) &< (m_0^\beta - u_0^\alpha - b)^2 + G(x_0^\alpha). \end{aligned} \quad (5)$$

In a fully revealing equilibrium, the encoder and the decoder policies are injective, thus these policies can be taken as identity functions; i.e., $x_0 = \gamma_0^e(m_0) = m_0$ and $u_0 = \gamma_0^d(x_0) = x_0 = m_0$. If we let $m_0^\alpha \rightarrow m_0^\beta$, then (5)

becomes

$$\begin{aligned} G(m_0^\alpha) - G(m_0^\beta) &< (m_0^\alpha - m_0^\beta - b)^2 - b^2 \rightarrow 0 \\ G(m_0^\beta) - G(m_0^\alpha) &< (m_0^\beta - m_0^\alpha - b)^2 - b^2 \rightarrow 0. \end{aligned} \quad (6)$$

Thus if $m_0^\alpha \rightarrow m_0^\beta$ we must have $G(m_0^\alpha) \rightarrow G(m_0^\beta)$ which implies that $G(x_0)$ is continuous at $x_0 = m_0^\beta$. Since this is valid for any m_0^β and $G(x_0)$ can take finitely many values, $G(x_0)$ cannot have any jumps. Thus, it can be deduced that $G(x_0)$ is a constant function which is equivalent to say that if the first stage equilibrium is fully revealing, then the second stage equilibrium is constructed independently from the first stage equilibrium. Then (6) reduces to $b^2 < (m_0^\alpha - m_0^\beta - b)^2$ and $b^2 < (m_0^\beta - m_0^\alpha - b)^2$. After simplifications, it can be found that these inequalities are satisfied simultaneously if $|m_0^\alpha - m_0^\beta| > 2|b|$; however, this contradicts with the $m_0^\alpha \rightarrow m_0^\beta$ assumption. Hence, the equilibrium cannot be fully informative at the first stage. ■

Now we can extend the 2-stage repeated game results to the N -stage repeated games as follows:

Corollary 3.1: Assuming the deterministic equilibrium for the N -stage repeated cheap talk game, the equilibrium policies for the final stage must be quantized almost surely for any collection of policies $(\gamma_0^e, \gamma_1^e, \dots, \gamma_{N-2}^e), (\gamma_0^d, \gamma_1^d, \dots, \gamma_{N-2}^d)$. If the sources m_0, m_1, \dots, m_{N-1} are uniform on $[0, 1]$, all stages except the final one cannot have fully revealing equilibria.

We note now two extensions of the results stated above, these will be reported in an extended paper.

Remark 3.1:

- i) The uniform source assumption can be relaxed for Theorem 3.2. Namely, any type of source that results in finitely many quantization bins in an equilibrium satisfies the statement; e.g., when the source admits an exponentially distributed real random variable (see [10]).
- ii) The results here are also applicable when the source is Markovian using similar arguments together with stochastic realization results for Markov sources.

B. Dynamic Cheap Talk under Stackelberg Equilibria

In this part, the cheap talk game is analyzed under the Stackelberg assumption; i.e., the encoder knows the policy of the decoder. In this case, admittedly the problem is less interesting.

Theorem 3.3: An equilibrium has to be fully revealing in the dynamic Stackelberg cheap talk game regardless of the source model.

Proof: We will use the properties of iterated expectations in the analysis. Recall that the total decoder cost is $J^d(\gamma^e, \gamma^d) = E \left[\sum_{k=0}^{N-1} (m_k - u_k)^2 \right]$. Considering the last stage, the goal of the decoder is to minimize $J_{N-1}^d(\gamma_{N-1}^e, \gamma_{N-1}^d) = E[(m_{N-1} - u_{N-1})^2 | \mathcal{I}_{N-1}^d]$ by choosing the optimal action $u_{N-1}^* = \gamma_{N-1}^{*,d}(\mathcal{I}_{N-1}^d) = E[m_{N-1} | \mathcal{I}_{N-1}^d]$. For the previous stage, the goal of the decoder is to minimize $J_{N-2}^d(\gamma_{N-2}^e, \gamma_{N-2}^d) = E[(m_{N-2} - u_{N-2})^2 + J_{N-1}^{*,d}(\gamma_{N-1}^e, \gamma_{N-1}^d) | \mathcal{I}_{N-2}^d]$ by choosing the optimal action $u_{N-2}^* = \gamma_{N-2}^{*,d}(\mathcal{I}_{N-2}^d)$. Since $J_{N-1}^{*,d}(\gamma_{N-1}^e, \gamma_{N-1}^d)$ is not affected by the choice of γ_{N-2}^d ,

the goal of the decoder is equivalent to the minimization of $E[(m_{N-2} - u_{N-2})^2 | \mathcal{I}_{N-2}^d]$ at this stage. Thus, the optimal policy is $u_{N-2}^* = \gamma_{N-2}^{*,d}(\mathcal{I}_{N-2}^d) = E[m_{N-2} | \mathcal{I}_{N-2}^d]$. Similarly, since the actions taken by the decoder do not affect the future states and encoder policies, the optimal decoder actions can be found as $u_k^* = \gamma_k^{*,d}(\mathcal{I}_k^d) = E[m_k | \mathcal{I}_k^d] = E[m_k | x_{[0,k]}]$ for $k = 0, 1, \dots, N-1$.

Due to the Stackelberg assumption, the encoder knows that the decoder will use $u_k^* = \gamma_k^{*,d}(\mathcal{I}_k^d) = E[m_k | \mathcal{I}_k^d]$ for each stage $k = 0, 1, \dots, N-1$. By using this assumption and the smoothing property of the expectation, the total encoder cost can be written as $J^e(\gamma^e, \gamma^d) = E\left[\sum_{k=0}^{N-1} (m_k - u_k - b)^2\right] = E\left[\sum_{k=0}^{N-1} (m_k - u_k)^2\right] + Nb^2$. Thus, as in the one-stage game setup [5, Theorem 3.4], the goals of the encoder and the decoder become essentially the same in the Stackelberg game setup, which effectively reduces the game setup to a team setup, resulting in fully informative equilibria; i.e. the encoder reveals all of its information. ■

IV. DYNAMIC QUADRATIC GAUSSIAN SIGNALING GAMES

The dynamic signaling game setup is similar to the dynamic cheap talk setup except that there exists an additive Gaussian noise channel between the encoder and decoder at each stage, and the encoder has a *soft* power constraint.

Here, the source is assumed to be a Markovian source with initial Gaussian distribution; i.e. $M_0 \sim \mathcal{N}(0, \sigma_{M_0}^2)$ and $M_{k+1} = aM_k + V_k$ where $a \in \mathbb{R}$ and $V_k \sim \mathcal{N}(0, \sigma_{V_k}^2)$ is an i.i.d. Gaussian noise sequence for $k = 0, 1, \dots, N-2$. The channels between the encoder and the decoder are assumed to be i.i.d. additive Gaussian channels; i.e. $W_k \sim \mathcal{N}(0, \sigma_{W_k}^2)$, and W_k and V_l are independent for $k = 0, 1, \dots, N-1$ and $l = 0, 1, \dots, N-2$.

In each stage of the game, say stage- k , the encoder aims to encode the \mathbb{R} -valued random variable M_k to the decoder by knowing the values of $\mathcal{I}_k^e = \{m_{[0,k]}, y_{[0,k-1]}\}$ with $\mathcal{I}_0^e = \{m_0\}$. Let X_k denote the \mathbb{R} -valued random variable which is transmitted to the decoder. During the transmission, the zero mean Gaussian noise with a variance of $\sigma_{W_k}^2$ is added to X_k ; hence, the decoder receives $Y_k = X_k + W_k$. The decoder, upon receiving Y_k , generates its optimal decision U_k which is also \mathbb{R} -valued by knowing the values of $\mathcal{I}_k^d = \{y_{[0,k]}\}$. We only consider the deterministic policies; i.e., $x_k = \gamma_k^e(\mathcal{I}_k^e)$ and $u_k = \gamma_k^d(\mathcal{I}_k^d)$.

The goal of the encoder is to find the optimal policy sequence $\gamma^e = \{\gamma_0^e, \gamma_1^e, \dots, \gamma_{N-1}^e\}$ that minimizes

$$J^e(\gamma^e, \gamma^d) = \sum_{k=0}^{N-1} \int c_k^e(m_k, u_k) P(dy_k | x_k) P(dm_k)$$

whereas the goal of the decoder is to find the optimal policy sequence $\gamma^d = \{\gamma_0^d, \gamma_1^d, \dots, \gamma_{N-1}^d\}$ that minimizes

$$J^d(\gamma^e, \gamma^d) = \sum_{k=0}^{N-1} \int c_k^d(m_k, u_k) P(dy_k | x_k) P(dm_k).$$

The cost functions are modified as $c_k^e(m_k, x_k, u_k) = (m_k - u_k - b)^2 + \lambda x_k^2$ and $c_k^d(m_k, u_k) = (m_k - u_k)^2$.

Note that a power constraint with an associated multiplier is appended to the cost function of the encoder, which corresponds to power limitation for transmitters in practice. If $\lambda = 0$, this corresponds to the setup with no power constraint at the encoder. For the purpose of illustration, the system model of the 2-stage dynamic signaling game is shown in Fig. 1-b.

A. Dynamic Signaling Game under Stackelberg Equilibria

In this part, the signaling game is analyzed under the Stackelberg assumption; i.e., the encoder knows the policy of the decoder.

Theorem 4.1: An equilibrium has to be always linear in the dynamic Stackelberg signaling game.

Proof: Similar to the dynamic Stackelberg cheap talk analysis in Theorem 3.3, the optimal decoder actions can be found as $u_k^* = \gamma_k^{*,d}(\mathcal{I}_k^d) = E[m_k | \mathcal{I}_k^d] = E[m_k | y_{[0,k]}]$ for $k = 0, 1, \dots, N-1$.

Due to the Stackelberg assumption, the encoder knows that the decoder will use $u_k^* = \gamma_k^{*,d}(\mathcal{I}_k^d) = E[m_k | \mathcal{I}_k^d]$ for each stage $k = 0, 1, \dots, N-1$. Based on this assumption and the smoothing property of the expectation, the total encoder cost can be written as $J^e(\gamma^e, \gamma^d) = E\left[\sum_{k=0}^{N-1} (m_k - u_k - b)^2 + \lambda x_k^2\right] = E\left[\sum_{k=0}^{N-1} E[(m_k - E[m_k | \mathcal{I}_k^d])^2 + b^2 + \lambda x_k^2 | \mathcal{I}_k^d]\right]$.

This problem is an instance of studied problems in [11, Chp.11] and [12], and can be reduced to a team problem where both the encoder and the decoder are minimizing the same expression.

From [12], the lower bound of $\sum_{k=0}^{N-1} \lambda_k P_k^2 + \inf_{\gamma^e, \gamma^d; E[x_k^2]=P_k^2, \forall k} E\left[\sum_{k=0}^{N-1} (a'_k(u_k - b'_k m_k)^2)\right]$ is

achieved when $u_k^* = \gamma_k^{*,d} = b'_k E[m_k | y_{[0,k]}]$ for $k = 0, 1, \dots, N-1$, and $x_0 = \gamma_0^{*,e}(m_0) = \eta_0 m_0$ and $x_k = \gamma_k^{*,e}(m_{[0,k]}, y_{[0,k-1]}) = \eta_k (m_k - E[m_k | y_{[0,k]}])$ for $k = 1, 2, \dots, N-1$. Here, η_k 's satisfy the recursion (for $k = 1, 2, \dots, N-1$) $\eta_k^2 = \frac{P_k^2}{a^2 \Delta_{k-1} + \sigma_V^2}$

with the initial condition $\eta_0^2 = \frac{P_0^2}{\sigma_{M_0}^2}$ and Δ_k 's satisfy the recursion (for $k = 1, 2, \dots, N-1$) $\Delta_k = \frac{\sigma_{W_k}^2}{P_k^2 + \sigma_{W_k}^2} (a^2 \Delta_{k-1} + \sigma_V^2)$ with the initial condition

$\Delta_0 = \frac{\sigma_{M_0}^2 \sigma_{W_0}^2}{P_0^2 + \sigma_{W_0}^2}$. Thus, the lower bound of $\sum_{k=0}^{N-1} \lambda_k P_k^2 +$

$\inf_{\gamma^e; E[x_k^2]=P_k^2, \forall k} E\left[\sum_{k=0}^{N-1} (a'_k b'_k (m_k - E[m_k | y_{[0,k]}])^2)\right]$ is achieved when $x_0 = \gamma_0^{*,e}(m_0) = \eta_0 m_0$ and $x_k = \gamma_k^{*,e}(m_{[0,k]}, y_{[0,k-1]}) = \eta_k (m_k - E[m_k | y_{[0,k]}])$ for $k = 1, 2, \dots, N-1$ which implies that the encoder also uses linear policies at each stage. Hence, the only equilibrium in the noisy dynamic Stackelberg setup of the signaling game is the linear equilibrium. ■

B. Nash Equilibria Analysis of N -stage Dynamic Signaling Games

In this section, for the N -stage dynamic signaling game, the optimality of an affine encoder is proved for an affine decoder, and the optimality of an affine decoder is shown for an affine encoder.

Theorem 4.2:

- i) If the decoder uses affine policies at all stages, then the encoder will also be affine at all stages.
- ii) If the encoder uses affine policies at all stages, then the decoder will also be affine at all stages.

Proof: Here, the proofs are presented for the 2-stage case. The approach here can be extended to the N -stage case.

- i) Let the decoder policies be $u_0 = \gamma_0^d(y_0) = Ky_0 + L$ and $u_1 = \gamma_1^d(y_0, y_1) = M_0y_0 + M_1y_1 + N$ where K, L, M_0, M_1 and N are scalars. With $y_1 = x_1 + w_1$, it follows that $u_1 = M_0y_0 + M_1x_1 + M_1w_1 + N$. Then, by completing the squares, the second stage encoder cost can be written as

$$\begin{aligned} J_1^{*,e} &= \min_{x_1=\gamma_1^e(m_1, y_0)} E \left[(m_1 - u_1 - b)^2 + \lambda x_1^2 \right] \\ &= \min_{\gamma_1^e(m_1, y_0)} E \left[\frac{\lambda}{M_1^2 + \lambda} (m_1 - M_0y_0 - N - b)^2 \right. \\ &\quad \left. + M_1^2 \sigma_{W_1}^2 + (M_1^2 + \lambda) \right. \\ &\quad \left. E \left[\left(x_1 - \frac{M_1(m_1 - M_0y_0 - N - b)}{M_1^2 + \lambda} \right)^2 \right] \right]. \end{aligned}$$

Hence, the optimal $\gamma_1^e(m_1, y_0)$ can be chosen as

$$\gamma_1^{*,e}(m_1, y_0) = \frac{M_1}{M_1^2 + \lambda} (m_1 - M_0y_0 - N - b).$$

Then the total cost of the encoder is the following:

$$J_0^{*,e} = \min_{x_0=\gamma_0^e(m_0)} E \left[(m_0 - u_0 - b)^2 + \lambda x_0^2 + J_1^{*,e} \right].$$

After the simplifications and completing the squares, the optimal first stage encoder policy can be found as

$$\begin{aligned} \gamma_0^{*,e}(m_0) &= \frac{KM_1^2 + \lambda K + \lambda a M_0}{(\lambda + K^2)(\lambda + M_1^2) + \lambda M_0^2} m_0 \\ &\quad - \frac{K(M_1^2 + \lambda)(L + b) + \lambda M_0(N + b)}{(\lambda + K^2)(\lambda + M_1^2) + \lambda M_0^2}. \end{aligned}$$

- ii) Since this result is immediate through MMSE properties for Gaussian variables, we have omitted the proof. ■

V. CONCLUDING REMARKS

In this study, dynamic and repeated quadratic cheap talk and signaling game problems are analyzed. For the cheap talk problem under Nash equilibria, we show that the last stage equilibria are quantized for any source with arbitrary distribution, and fully revealing equilibria cannot exist for some source models (see Remark 3.1); whereas, for the dynamic Stackelberg cheap talk, the equilibria must be fully revealing regardless of the source model. In the dynamic signaling game where the transmission of a Gaussian source over a Gaussian channel is considered, for scalar sources under Stackelberg equilibria, the only equilibrium is the linear equilibrium; while, for the Nash equilibria, affine policies constitute an invariant subspace under best response maps.

REFERENCES

- [1] V. P. Crawford and J. Sobel, "Strategic information transmission," *Econometrica*, vol. 50, pp. 1431–1451, 1982.
- [2] V. Poor, *An Introduction to Signal Detection and Estimation*. Springer, 1994.
- [3] T. Başar and G. Olsder, *Dynamic Noncooperative Game Theory*. Philadelphia, PA: SIAM Classics in Applied Mathematics, 1999.
- [4] I. Shames, A. Teixeira, H. Sandberg, and K. Johansson, "Agents misbehaving in a network: a vice or a virtue?" *IEEE Network*, vol. 26, no. 3, pp. 35–40, May 2012.
- [5] S. Sarıtaş, S. Yüksel, and S. Gezici, "Quadratic multi-dimensional signaling games and affine equilibria," *IEEE Transactions on Automatic Control*, 2017, to appear. [Online]. Available: <http://arxiv.org/abs/1503.04360>
- [6] F. Farokhi, A. M. H. Teixeira, and C. Langbort, "Estimation with strategic sensors." 2014. [Online]. Available: <http://arxiv.org/abs/1402.4031>
- [7] E. Akyol, C. Langbort, and T. Başar, "Strategic compression and transmission of information," in *IEEE Information Theory Workshop*, 2015.
- [8] M. Golosov, V. Skreta, A. Tsyvinski, and A. Wilson, "Dynamic strategic information transmission," *Journal of Economic Theory*, vol. 151, pp. 304–341, 2014.
- [9] J. Renault, E. Solan, and N. Vieille, "Dynamic sender receiver games," *Journal of Economic Theory*, vol. 148, no. 2, pp. 502–534, 2013.
- [10] S. Fabricius, P. Furrer, S. Kerner, T. Linder, and S. Yüksel, "Game theory and information, Queen's University, MTHE 493 Technical Report," Apr. 2014.
- [11] S. Yüksel and T. Başar, *Stochastic Networked Control Systems: Stabilization and Optimization under Information Constraints*. Boston, MA: Birkhäuser, 2013.
- [12] R. Bansal and T. Başar, "Simultaneous design of measurement and control strategies for stochastic systems with feedback," *Automatica*, vol. 25, no. 5, pp. 679–694, 1989.