Dynamic signaling games with quadratic criteria under Nash and Stackelberg equilibria☆

Serkan Sarıtaş [a,*], Serdar Yüksel [b], Sinan Gezici [c]

[a] *Division of Decision and Control Systems, KTH Royal Institute of Technology, SE 10044, Stockholm, Sweden*
[b] *Department of Mathematics and Statistics, Queen's University, K7L 3N6, Kingston, Ontario, Canada*
[c] *Department of Electrical and Electronics Engineering, Bilkent University, 06800, Ankara, Turkey*

ARTICLE INFO

ABSTRACT

This paper considers dynamic (multi-stage) signaling games involving an encoder and a decoder who have subjective models on the cost functions. We consider both Nash (simultaneous-move) and Stackelberg (leader–follower) equilibria of dynamic signaling games under quadratic criteria. For the multi-stage scalar cheap talk, we show that the final stage equilibrium is always quantized and under further conditions the equilibria for all time stages must be quantized. In contrast, the Stackelberg equilibria are always fully revealing. In the multi-stage signaling game where the transmission of a Gauss–Markov source over a memoryless Gaussian channel is considered, affine policies constitute an invariant subspace under best response maps for Nash equilibria; whereas the Stackelberg equilibria always admit linear policies for scalar sources but such policies may be non-linear for multi-dimensional sources. We obtain an explicit recursion for optimal linear encoding policies for multi-dimensional sources, and derive conditions under which Stackelberg equilibria are informative.

© 2020 Elsevier Ltd. All rights reserved.

1. Introduction

Signaling games and cheap talk are concerned with a class of Bayesian games where an informed player (encoder or sender) transmits information to another player (decoder or receiver). In these problems, the objective functions of the players are not aligned unlike the ones in the classical communication problems. The single-stage cheap talk problem was studied by Crawford and Sobel (1982), who obtained the surprising result that under some technical conditions on the cost functions, the cheap talk problem only admits equilibria that involve quantized encoding policies. This is in contrast to the usual communication/information theoretic case where the goals are aligned. In this paper, we consider multi-stage signaling games. The details are presented in the following.

A single-stage (static) *cheap talk* problem can be formulated as follows: An informed player (encoder) knows the value of the $\mathbb{M}$-valued random variable M and transmits the $\mathbb{X}$-valued random variable X to another player (decoder), who generates his $\mathbb{M}$-valued optimal decision U upon receiving X. The policies of the encoder and the decoder are assumed to be deterministic; i.e., $x = \gamma^e(m)$ and $u = \gamma^d(x) = \gamma^d(\gamma^e(m))$. Let $c^e(m, u)$ and $c^d(m, u)$ denote the cost functions of the encoder and the decoder, respectively, when the action u is taken for the corresponding message m. Then, given the encoding and decoding policies, the encoder's induced expected cost is $J^e\left(\gamma^e, \gamma^d\right) = \mathbb{E}[c^e(m, u)]$, whereas, the decoder's induced expected cost is $J^d\left(\gamma^e, \gamma^d\right) = \mathbb{E}\left[c^d(m, u)\right]$. If the transmitted signal x is also an explicit part of the cost functions c^e and/or c^d, then the communication between the players is not costless and the formulation turns into a *signaling game* problem. Such problems are studied under the tools and concepts provided by *game theory* since the goals are not aligned. Although the encoder and the decoder act sequentially in the game as described above, how and when the decisions are made and the nature of the commitments to the announced policies significantly affect the analysis of the equilibrium structure. Here, two different types of equilibria are investigated: the *Nash equilibrium*, in which the encoder and the decoder make simultaneous decisions, and the *Stackelberg equilibrium*, in which the encoder and the decoder make sequential decisions where the

encoder is the leader and the decoder is the follower.[1] In this paper, the terms *Nash game* and the *simultaneous-move game*[2] will be used interchangeably, and similarly, the *Stackelberg game* and the *leader–follower game* will be used interchangeably.

In the simultaneous-move game, the encoder and the decoder announce their policies at the same time, and a pair of policies $(\gamma^{*,e}, \gamma^{*,d})$ is said to be a **Nash equilibrium** (Başar & Olsder, 1999) if

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

where Γ^e and Γ^d are the sets of all deterministic (and Borel measurable) functions from $\mathbb{M}$ to $\mathbb{X}$ and from $\mathbb{X}$ to $\mathbb{M}$, respectively. As observed from the definition (1), under the Nash equilibrium, each individual player chooses an optimal strategy given the strategies chosen by the other players.

On the other hand, in a leader–follower game, the leader (encoder) commits to and announces his optimal policy before the follower (decoder) does, the follower observes what the leader is committed to before choosing and announcing his optimal policy, and a pair of policies $(\gamma^{*,e}, \gamma^{*,d})$ is a **Stackelberg equilibrium** (Başar & Olsder, 1999) if

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

As it can be seen from the definition (2), the decoder takes his optimal action $\gamma^{*,d}(\gamma^e)$ after observing the policy of the encoder γ^e. In the Stackelberg game, the leader cannot backtrack on his commitment, but has a leadership role since he can manipulate the follower by anticipating follower's actions; which corresponds to Bayesian persuasion problems in economics (Kamenica & Gentzkow, 2011). We provide a further literature review on this later.

If an equilibrium is achieved when $\gamma^{*,e}$ is non-informative (e.g., the transmitted message and the source are independent) and $\gamma^{*,d}$ uses only the prior information (since the received message is useless), then we call such an equilibrium a *non-informative (babbling) equilibrium*, which always exists for cheap talk (Crawford & Sobel, 1982).

Heretofore, only *single-stage games* are considered. If a game is played over a number of time periods, the game is called a *multi-stage game*. In this paper, with the term *dynamic*, we will refer to *multi-stage game* setups, which also has been the usage in the prior literature (Golosov, Skreta, Tsyvinski, & Wilson, 2014); even though strictly speaking a single stage setup may also be viewed to be dynamic (Yüksel & Başar, 2013) since the information available to the decoder is totally determined by encoder's actions.

Let $m_{[0,N-1]} = \{m_0, m_1, \ldots, m_{N-1}\}$ be a collection of random variables to be encoded sequentially (causally) to a decoder. At the kth stage of an N-stage game, the encoder knows $\mathcal{I}_k^e = \{m_{[0,k]}, x_{[0,k-1]}\}$ with $\mathcal{I}_0^e = \{m_0\}$, and transmits x_k to the decoder who generates his optimal decision by knowing $\mathcal{I}_k^d = \{x_{[0,k]}\}$. 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 encoder's goal is to minimize

$$
J^e\left(\gamma_{[0,N-1]}^e, \gamma_{[0,N-1]}^d\right) = \mathbb{E}\left[\sum_{k=0}^{N-1} c_k^e(m_k, u_k)\right],
\tag{3}
$$

whereas the decoder's goal is to minimize

$$
J^d\left(\gamma_{[0,N-1]}^e, \gamma_{[0,N-1]}^d\right) = \mathbb{E}\left[\sum_{k=0}^{N-1} c_k^d(m_k, u_k)\right]
\tag{4}
$$

by finding the optimal policy sequences $\gamma_{[0,N-1]}^{*,e} = \{\gamma_0^{*,e}, \gamma_1^{*,e}, \ldots, \gamma_{N-1}^{*,e}\}$ and $\gamma_{[0,N-1]}^{*,d} = \{\gamma_0^{*,d}, \gamma_1^{*,d}, \ldots, \gamma_{N-1}^{*,d}\}$, respectively. Using the encoder cost in (3) and the decoder cost in (4), the Nash equilibrium and the Stackelberg equilibrium for multi-stage games can be defined similarly 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; thus the equilibria are to be interpreted under a *perfect Bayesian equilibria* concept (Fudenberg & Tirole, 1991). Since we assume such a (perfect Bayesian) framework, the equilibria lead to sub-game perfection and each decision maker performs optimal Bayesian decisions for every realized play path. For example, more general Nash equilibrium scenarios such as non-credible threats or equilibria that are not strong time-consistent (Yüksel & Başar, 2013, Definition 2.4.1) may not be considered. We also note that both Nash and Stackelberg equilibrium concepts find various applications depending on the assumptions on the leader, that is, the encoder, in view of the commitment conditions (Korzhyk, Yin, Kiekintveld, Conitzer, & Tambe, 2011).

In this paper, 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 Crawford and Sobel (1982) and Sarıtaş, Yüksel, and Gezici (2017b).

1.1. Literature review

As noted in Sarıtaş et al. (2017b), for team problems, although it is difficult to obtain optimal solutions under general information structures, it is apparent that more information provided to any of the decision makers does not negatively affect the utility of the players. There is also a well-defined partial order of information structures as studied by Blackwell (1951) and Yüksel and Başar (2013). However, for general non-zero-sum game problems, informational aspects are very challenging to address; more information can have negative effects on some or even all of the players in a system, see e.g. Hirshleifer (1971).

The cheap talk and signaling game problems are applicable in networked control systems when a communication channel exists among competitive and non-cooperative decision makers. For example, in a smart grid application, there may be strategic sensors in the system (Shames, Teixeira, Sandberg, & Johansson, 2012) that wish to change the equilibrium for their own interests through reporting incorrect measurement values, e.g., to enhance prolonged use in the system. For further applications, see Larrousse, Beaude, and Lasaulce (2014) and Shames et al. (2012). All of these applications lead to a new framework where the value of information and its utilization have a drastic impact on the system under consideration.

The reader is referred to Sarıtaş et al. (2017b) for further discussion and references on single-stage signaling games. On the multi-stage side, much of the literature has focused on Stackelberg equilibria as we note below. A notable exception is Golosov et al. (2014), where the multi-stage extension of the setup of Crawford and Sobel is analyzed for a source which is a fixed random variable distributed according to some density on [0, 1] (see Theorem 8 for a detailed discussion on this very relevant paper). For the case in which the non-alignment between the cost functions of the encoder and the decoder (i.e., the bias term b) is a function of a Gaussian random variable (r.v.) correlated

---

[1] For the Nash equilibrium, the encoder and the decoder take actions sequentially but do not announce their policies beforehand (i.e., they announce simultaneously) while they do so for the Stackelberg equilibrium.

[2] Note that throughout the manuscript, *simultaneous-move* refers to *simultaneous-announcement*, see[1].

with the Gaussian source and secret to the decoder (contrarily to the original case in which it is fixed and known to the decoder (Crawford & Sobel, 1982), which is also studied in Sarıtaş et al. (2017b) and in this manuscript),[3] the Stackelberg equilibrium under quadratic costs is investigated in Akyol et al. (2017), Farokhi et al. (2017) and Sayin et al. (2019). The Stackelberg equilibrium is characterized under the noiseless setting and the affine transmitter and receiver assumptions in Farokhi et al. (2017), whereas the optimality of linear sender and receiver policies is proved in Akyol et al. (2017) at the unique equilibrium under both noiseless and Gaussian noise settings without any linearity assumptions. An information theoretic formulation of the Bayesian persuasion problem (Kamenica & Gentzkow, 2011) is studied in Akyol et al. (2017) for general (not necessarily Gaussian) sources, and recently also in Treust and Tomala (2019) and Treust and Tomala (2018) with finite state and action spaces by assuming an exogenous noise (limited channel capacity) and side information, respectively. The Stackelberg equilibrium of the multi-stage multi-dimensional Gaussian signaling game under general quadratic cost functions is investigated in Sayin et al. (2019) and the joint optimality of linear encoder and decoder strategies is proved in noiseless settings under a finite horizon when the private state of the encoder is a controlled Gauss–Markov process. Vasal and Anastasopoulos (2016) investigate the Nash equilibrium of a multi-stage linear quadratic Gaussian game with asymmetric information, and it is shown that under certain conditions, players' strategies are linear in their private types. The cheap talk with finite state and action spaces and multiple round of pre-play communication is investigated where both the encoder and the decoder take costly actions at the end of the pre-play communication in Aumann and Hart (2003), and it is proved that the multiple round of communication improves information revelation. The dynamic extension to the optimal information disclosure (Kamenica & Gentzkow, 2011) with a finite state space and a finite number of periods is considered in Ely, Frankel, and Kamenica (2015), in which the sender commits to a policy similar to the Stackelberg case. In our earlier work (Sarıtaş et al., 2017b), we considered both Nash equilibria and Stackelberg equilibria. In this paper, we build on Sarıtaş et al. (2017b), and extend the analysis to the multi-stage case. We also note our related conference papers (Sarıtaş, Yüksel, & Gezici, 2016) and (Sarıtaş, Yüksel, & Gezici, 2017a); compared to these, in this paper we have significantly enhanced the analysis both by including complete proofs and more general results.

1.2. Contributions

(i) We show that in the multi-stage cheap talk game under Nash equilibria, the last stage equilibria are quantized for scalar i.i.d. and Markov sources with arbitrary conditional probability measures and not fully revealing for multi-dimensional sources, whereas the equilibrium must be fully revealing in the multi-stage cheap talk game under Stackelberg equilibria for both scalar and multi-dimensional sources. Further, for scalar i.i.d. sources, the quantized nature of the Nash equilibrium for all stages is established under mild conditions.

(ii) For the multi-stage signaling game under Nash equilibria, it is shown that affine encoder and decoder policies constitute an invariant subspace under best response dynamics. We

(a) 2-stage cheap talk.
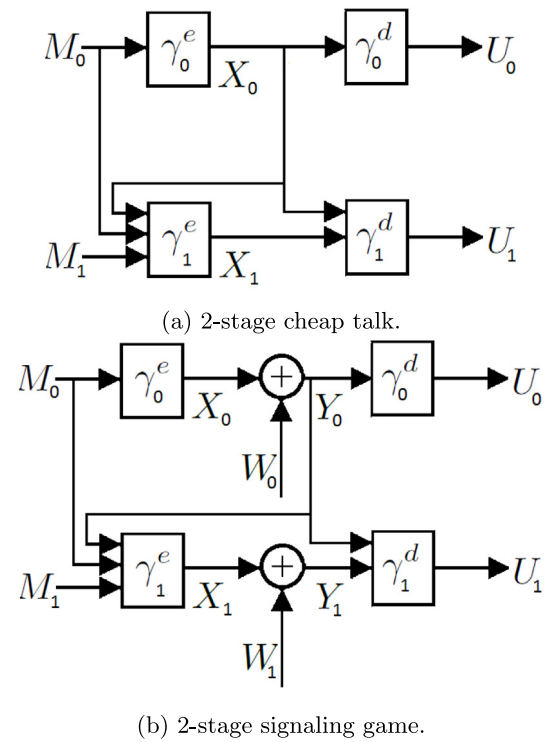

(b) 2-stage signaling game.

Fig. 1. Signaling game models.

provide conditions for the existence of informative Stackelberg equilibria for scalar Gauss–Markov sources and scalar Gaussian channels where we also show that Stackelberg equilibria are always linear for scalar sources and channels, which is not always the case for multi-dimensional setups. For multi-dimensional setups, a dynamic programming formulation is presented for Stackelberg equilibria when the encoders are linear.

2. Multi-stage cheap talk

For the purpose of illustration, the system model of the 2-stage cheap talk is depicted in Fig. 1(a). Before presenting the technical results related to the multi-stage cheap talk, we provide the result on the static (single-stage) cheap talk with deterministic encoders from our previous study, which will be utilized in several results later on this paper.

Theorem 1 (*Sarıtaş et al., 2017b, Theorem 3.2*). *Let the strategy set of the encoder consist of the set of all measurable (deterministic) functions from $\mathbb{M}$ to $\mathbb{X}$. Then, an equilibrium encoder policy has to be quantized almost surely, that is, it is equivalent to a quantized policy for the encoder in the sense that the performance of any equilibrium encoder policy is equivalent to the performance of a quantized encoder policy. Furthermore, the quantization bins are convex.*

To facilitate our analysis to handle certain intricacies that arise due to the multi-stage setup in this paper, in the following, we state that the result in Theorem 1 also holds when the encoder is allowed to adapt randomized encoding policies by extending (Crawford & Sobel, 1982, Lemma 1) as follows:

Theorem 2. *The conclusion of Theorem 1 holds if the policy space of the encoder is extended to the set of all stochastic kernels from*

$\mathbb{M}$ to $\mathbb{X}$ for any arbitrary source.[4] *That is, even when the encoder is allowed to use private randomization, all equilibria are equivalent to those that are attained by quantized equilibria.*

Proof (*Crawford & Sobel, 1982, Lemma 1*)**.** proves that all equilibria with a randomized encoder have finitely many partitions when the source has bounded support. Theorem 1 extends this result to a countable number of partitions (i.e., distinct decoder actions must differ by at least $2|b|$) for any source with an arbitrary probability measure, but assuming a deterministic encoder. The proof methods in both of Crawford and Sobel (1982, Lemma 1) and Theorem 1 can be combined, which implies that equilibria must be quantized. □

Theorem 2 will be used crucially in the following analysis; since in a multi-stage game, at a given time stage, the source variables from the earlier stages can serve as private randomness for the encoder. As a prelude to the more general Markov source setup, we first analyze the multi-stage cheap talk game with an i.i.d. scalar source.

2.1. Multi-stage game with an i.i.d. source

Theorem 3. *In the N-stage repeated cheap talk game, the equilibrium policies for the final stage encoder must be quantized for any collection of policies $\left(\gamma^e_{[0,N-2]}, \gamma^d_{[0,N-2]}\right)$ and for any real-valued source model with arbitrary probability measure $P(\mathrm{d}m_{N-1})$.*

Proof. Here, we prove the result for the 2-stage setup, the extension to multiple stages is merely technical, as we comment on at the end of the proof. Let $c^e_1(m_1, u_1)$ be the second stage cost function of the encoder. Then the expected cost of the second stage encoder J^e_1 is

$$
\begin{aligned}
J^e_1 &= \int P(\mathrm{d}m_0, \mathrm{d}m_1, \mathrm{d}x_0, \mathrm{d}x_1)\, c^e_1(m_1, u_1) \\
&= \int P(\mathrm{d}x_0) \int P(\mathrm{d}m_1|x_0)\, P(\mathrm{d}m_0|m_1, x_0) \\
&\quad \times c^e_1(m_1, \gamma^d_1(x_0, \gamma^e_1(m_0, m_1, x_0))).
\end{aligned}
\tag{5}
$$

Since $P(\mathrm{d}m_1|x_0) = P(\mathrm{d}m_1)$ and $P(\mathrm{d}m_0|m_1, x_0) = P(\mathrm{d}m_0|x_0)$ for an i.i.d. source, the inner integral of (5) can be considered as an expression for a given x_0. Thus, given the second stage encoder and decoder policies $\gamma^e_1(m_0, m_1, x_0)$ and $\gamma^d_1(x_0, x_1)$, it is possible to define policies which are parametrized by the common information x_0 almost surely so that $\widehat{\gamma}^e_{x_0}(m_0, m_1) \triangleq \gamma^e_1(m_0, m_1, x_0)$ and $\widehat{\gamma}^d_{x_0}(x_1) \triangleq \gamma^d_1(x_0, x_1)$.

Now fix the first stage policies γ^e_0 and γ^d_0. Suppose that the second stage encoder does not use m_0; i.e., $\widehat{\gamma}^{e\prime}_{x_0}(m_1)$ is the policy of the second stage encoder. For the policies $\widehat{\gamma}^{e\prime}_{x_0}(m_1)$ and $\widehat{\gamma}^d_{x_0}(x_1)$, by using the second stage encoder cost function $H_{x_0}(m_1, u_1) \triangleq \mathbb{E}[(m_1 - u_1 - b)^2|x_0]$ and the bin arguments from Theorem 1, it can be deduced that, due to the continuity of $H_{x_0}(m_1, u_1)$ in m_1, the equilibrium policies for the second stage must be quantized for any collection of policies (γ^e_0, γ^d_0) and for any given x_0. Now let the second stage encoder use m_0; i.e., $\widehat{\gamma}^e_{x_0}(m_0, m_1)$ is the deterministic policy of the second stage encoder, which can be regarded 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 the source, m_1. From Theorem 2, the equilibrium is achievable with an encoder policy which uses

only m_1; i.e., $\widehat{\gamma}^{e*}_{x_0}(m_1)$ is an encoder policy at the equilibrium and thus the equilibria are quantized.

For the N-stage game, the common information of the final stage encoder and decoder becomes $x_{[0,N-2]}$, and $m_{[0,N-2]}$ is a vector valued random variable independent of the final stage source m_{N-1}. □

Assumption 4. The source m_k is so that the single-stage cheap-talk game satisfies the following:
 (i) There exists a finite upper bound on the number of quantization bins that any equilibrium admits.
 (ii) There exist finitely many equilibria corresponding to a given number of quantization bins.

A number of comments on Assumption 4 are in order: This assumption is not unrealistic, e.g., all sources with bounded support, and sources with sufficiently light tail such as the exponential distribution satisfy this property, provided that their associated densities are one-sided and the sign of b is negative (Sarıtaş, Furrer, Gezici, Linder, & Yüksel, 2019). A sufficient condition for Assumption 4 is that the source admits a bounded support (which would require by Crawford and Sobel (1982, Lemma 1) that there exists an upper bound on the number of bins in any equilibrium), and that a monotonicity condition (Crawford & Sobel, 1982, condition (M), or equivalently (M′)) holds, which characterize the behavior of equilibrium policies. Note though that this condition is much more than what is needed in Assumption 4, since it entails the uniqueness of equilibria for a given number of bins: The uniqueness of equilibria even for team problems with $b = 0$ (which amounts to uniqueness of fixed point solutions to the celebrated Lloyd–Max optimality conditions) holds under additional assumptions, such as log-concavity conditions (Kieffer, 1993, p. 1475), (Fleischer, 1964).

Theorem 5. *Under Assumption 4, all stages must have quantized equilibria with finitely many bins in the N-stage repeated cheap talk game.*

Proof. Consider first the 2-stage setup; i.e., given that the second stage has a quantized equilibrium by Theorem 3, the quantized nature of the first stage will be established. 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 (since m_0 is a random variable independent of the source m_1 as shown in Theorem 3), $F(m_0, x_0)$ can be written as $F(m_0, x_0) = (m_0 - \gamma^d_0(x_0) - b)^2 + G(x_0)$ where $G(x_0) \triangleq \mathbb{E}_{m_1}\left[\left(m_1 - \gamma^{*,d}_1(x_0, \gamma^{*,e}_1(m_1, x_0)) - b\right)^2 \Big| x_0\right]$ is the expected cost of the second stage encoder, and $\gamma^{*,e}_1$ and $\gamma^{*,d}_1$ are the second stage encoder and decoder policies at the equilibrium, respectively. Note that the second stage encoder cost $G(x_0)$ can take finitely many different values by Assumption 4(ii). Now define the equivalence classes T_{x_p} for every $x_p \in \mathbb{X}$ as $T_{x_p} = \{x \in \mathbb{X} : G(x) = G(x_p)\}$; i.e., the equivalence classes T_{x_p} keep the first stage encoder actions that result in the same second stage cost in the same set. Note that there are finitely many equivalence classes T_{x_p} since $G(x_0)$ can take finitely many different values.

If the number of bins of the first stage equilibrium is less than or equal to the number of the equivalence classes T_{x_p}, then the proof is complete; i.e., the first stage equilibrium is already quantized with finitely may bins. Otherwise, one of the equivalence classes T_{x_p} has at least two elements; say x^α_0 and x^β_0, which implies $G(x^\alpha_0) = G(x^\beta_0)$. Let corresponding bins of the actions x^α_0 and x^β_0 be $\mathcal{B}^\alpha_0$ and $\mathcal{B}^\beta_0$, respectively. Also let m^α_0 and m^β_0 represent any point in $\mathcal{B}^\alpha_0$ and $\mathcal{B}^\beta_0$, respectively; i.e., $m^\alpha_0 \in \mathcal{B}^\alpha_0$ and $m^\beta_0 \in \mathcal{B}^\beta_0$.

---

[4] P is a stochastic kernel from $\mathbb{M}$ to $\mathbb{X}$ if $P(\cdot|m)$ is a probability measure on $\mathcal{B}(\mathbb{X})$ for every $m \in \mathbb{M}$, and $P(A|\cdot)$ is a Borel measurable function of m for every $A \in \mathcal{B}(\mathbb{X})$.

The decoder chooses an action $u_0^\alpha = \gamma_0^d(x_0^\alpha)$ when the encoder sends $x_0^\alpha = \gamma_0^e(m_0^\alpha)$, and an 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 his total cost; further, we can assume that $u_0^\alpha < u_0^\beta$ without loss of generality. 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 $(u_0^\alpha - u_0^\beta)(u_0^\alpha + u_0^\beta - 2(m_0^\alpha - b)) < 0$ and $(u_0^\alpha - u_0^\beta)(u_0^\alpha + u_0^\beta - 2(m_0^\beta - b)) < 0$. Thus, we have $u_0^\alpha + u_0^\beta - 2(m_0^\alpha - b) > 0$ and $u_0^\alpha + u_0^\beta - 2(m_0^\beta - b) < 0$, that reduce to $m_0^\alpha < \frac{u_0^\alpha + u_0^\beta}{2} + b < m_0^\beta$. Since $u^\alpha = \mathbb{E}[m | m \in \mathcal{B}^\alpha]$ and $u^\beta = \mathbb{E}[m | m \in \mathcal{B}^\beta]$ at the equilibrium, $u_0^\alpha < \frac{u_0^\alpha + u_0^\beta}{2} + b < u_0^\beta \Rightarrow u_0^\beta - u_0^\alpha > 2|b|$ is obtained. Hence, there must be at least $2|b|$ distance between the actions of the first stage decoder which are in the same equivalence class. Therefore, the cardinality of any equivalence class T_{x_p} is finite due to Assumption 4(i). Further, there are finitely many equivalence classes T_{x_p} as shown above. These two results imply that the first stage equilibrium must be quantized with finitely many bins. Thus, due to Assumption 4(ii), there are finitely many equilibria in the first stage; i.e., the first stage encoder cost can take finitely many values.

For the N-stage game, we apply the similar recursion from the final stage to the first stage. It is already proven that the last two stage encoder cost can take finitely many values; thus, the same methods can be applied to show the quantized structure (with finitely many bins) of the equilibria for all stages. □

Remark 6. It is important to note that the first stage encoder minimizes his expected cost $J_0^e = \mathbb{E}[F(m_0, x_0)]$ by minimizing his cost $F(m_0, x_0)$ for every realizable m_0; this property will be later used as well.

2.2. Multi-stage game with a Markov source: Nash equilibria

Here, the source M_k is assumed to be real valued Markovian for $k \leq N - 1$. The following result generalizes Theorem 3, which only considered i.i.d. sources.

Theorem 7. *In the N-stage cheap talk game with a Markov source, the equilibrium policies for the final stage encoder must be quantized for any collection of policies $(\gamma_{[0,N-2]}^e, \gamma_{[0,N-2]}^d)$ and for any real-valued source model with arbitrary probability measure.*

Proof. Here, we prove the results for the 2-stage games as the extension is merely technical. Similar to that in Theorem 3, the expected cost of the second stage encoder can be written as (5). After following similar arguments, the second stage encoder policy becomes $\widehat{\gamma}_{x_0}^e(m_0, m_1) \stackrel{(a)}{=} \widehat{\gamma}_{x_0}^e(g(m_1, r), m_1) = \widetilde{\gamma}_{x_0}^e(m_1, r)$ where (a) holds since any stochastic kernel from a complete, separable and metric space to another one, $P(dm_0 | m_1)$, can be realized by some measurable function $m_0 = g(m_1, r)$ where r is a $[0, 1]$-valued independent random variable (see Gihman & Skorohod, 1979, Lemma 1.2, or in Borkar, 1993, Lemma 3.1). Hence, the equilibria are quantized by Theorem 3. □

As it can be observed from Theorem 5, to be able to claim that the equilibria for all stages are quantized, we require very strong conditions. In fact, in the absence of such conditions, the equilibria for a Markov source can be quite counterintuitive and even fully revealing as we observe in the following theorem due to Golosov et al. (2014).

Theorem 8 (*Golosov et al., 2014*)**.** *For a Markov source, there exist multi-stage cheap talk games with fully revealing equilibria.*

An example is presented in Golosov et al. (2014), where an individual source is transmitted repeatedly (thus the Markov source is a constant source) for a sufficiently small bias value. For such a source, the terminal stage conditional measure can be made atomic via a careful construction of equilibrium policies for earlier time stages; i.e., the defined separable groups/types and discrete/quantized stage-wise equilibria through multiple stages can lead to a fully informative equilibrium for the complete game.

Next, Nash equilibria of the multi-stage multi-dimensional cheap talk are analyzed. Since there may be discrete, non-discrete or even linear Nash equilibria in the single-stage multi-dimensional cheap talk by Sarıtaş et al. (2017b, Theorem 3.4), the equilibrium policies are more difficult to characterize; however, we state the following:

Theorem 9.

(i) *The Nash equilibrium cannot be fully revealing in the static (single-stage) multi-dimensional cheap talk when the source has positive measure for every non-empty open set.*

(ii) *The final stage Nash equilibria cannot be fully revealing in the multi-stage multi-dimensional cheap talk for i.i.d. and Markov sources when the conditional distribution $P(d\mathbf{m}_{N-1} | \mathbf{m}_{N-2})$ has positive measure for every non-empty open set.*

Proof. (i) Similar to the single-stage scalar case Theorem 1, in an equilibrium, define two cells C^α and C^β, any points in those cells as $\mathbf{m}^\alpha \in C^\alpha$ and $\mathbf{m}^\beta \in C^\beta$, and the actions of the decoder as $\mathbf{u}^\alpha$ and $\mathbf{u}^\beta$ when the encoder transmits $\mathbf{m}^\alpha$ and $\mathbf{m}^\beta$, respectively. Let $F(\mathbf{m}, \mathbf{u}) \triangleq \|\mathbf{m} - \mathbf{u} - \mathbf{b}\|^2$. Due to the equilibrium definitions from the view of the encoder; $F(\mathbf{m}^\alpha, \mathbf{u}^\alpha) < F(\mathbf{m}^\alpha, \mathbf{u}^\beta)$ and $F(\mathbf{m}^\beta, \mathbf{u}^\beta) < F(\mathbf{m}^\beta, \mathbf{u}^\alpha)$. Hence, there exists a hyperplane defined by $F(\mathbf{z}, \mathbf{u}^\alpha) = F(\mathbf{z}, \mathbf{u}^\beta) \Rightarrow \|(\mathbf{z}-\mathbf{b})-\mathbf{u}^\alpha\|^2 = \|(\mathbf{z}-\mathbf{b})-\mathbf{u}^\beta\|^2$. The hyperplane defined by the points $\mathbf{z}$ divides the space into two subspaces: let Z^α that contains $\mathbf{u}^\alpha$ and Z^β that contains $\mathbf{u}^\beta$ be those subspaces. $\mathcal{C}^\beta$ and Z^α are disjoint subspaces since $F(\mathbf{z} + \delta(\mathbf{u}^\beta - \mathbf{u}^\alpha), \mathbf{u}^\alpha) \geq F(\mathbf{z} + \delta(\mathbf{u}^\beta - \mathbf{u}^\alpha), \mathbf{u}^\beta)$ for any $\delta > 0$. Similarly, $\mathcal{C}^\alpha$ and Z^β are disjoint subspaces, too. Thus, the hyperplane defined by the points $\mathbf{z}$ must lie between $\mathbf{u}^\alpha$ and $\mathbf{u}^\beta$ which implies that the length of $\mathbf{b}$ along the $\mathbf{d} \triangleq \mathbf{u}^\beta - \mathbf{u}^\alpha$ direction should not exceed half of the distance between $\mathbf{u}^\alpha$ and $\mathbf{u}^\beta$; i.e., $\|\mathbf{b_d}\| \leq \|\mathbf{d}\|/2$, where $\mathbf{b_d}$ is the projection of $\mathbf{b}$ along the direction of $\mathbf{d}$. Since $\mathbf{d}$ can be any vector at a fully revealing equilibrium by the assumption on the source (i.e., the source has positive measure for every non-empty open set), $\|\mathbf{b_d}\| \leq \|\mathbf{d}\|/2$ cannot be satisfied unless $\mathbf{b} = \mathbf{0}$. Thus, there cannot be a fully revealing equilibrium in the static multi-dimensional cheap talk.

(ii) The proof is the multi-dimensional extension of Theorem 3 for i.i.d. sources, and Theorem 7 for Markov sources. □

2.3. Multi-stage cheap talk under Stackelberg equilibria

Here, the cheap talk game is analyzed under the Stackelberg formulation for both scalar and multi-dimensional sources. In this case, admittedly the problem is less interesting.

Theorem 10. *An equilibrium has to be fully revealing in the multi-stage Stackelberg cheap talk game regardless of the source model.*

Proof. The last stage decoder cost $J_{N-1}^d(\gamma_{N-1}^e, \gamma_{N-1}^d) = \mathbb{E}[\|\mathbf{m}_{N-1} - \mathbf{u}_{N-1}\|^2 | \mathcal{I}_{N-1}^d]$ is minimized by choosing the optimal action $\mathbf{u}_{N-1}^* = \gamma_{N-1}^{*,d}(\mathcal{I}_{N-1}^d) = \mathbb{E}[\mathbf{m}_{N-1} | \mathcal{I}_{N-1}^d]$. For the previous stage, the decoder can minimize $J_{N-2}^d(\gamma_{N-1}^{*,e}, \gamma_{N-2}^e, \gamma_{N-1}^{*,d}, \gamma_{N-2}^d) = \mathbb{E}[\|\mathbf{m}_{N-2} - \mathbf{u}_{N-2}\|^2 + J_{N-1}^{*,d}(\gamma_{N-1}^{*,e}, \gamma_{N-1}^{*,d}) | \mathcal{I}_{N-2}^d]$ by choosing his policy as $\mathbf{u}_{N-2}^* =$

$\gamma_{N-2}^{*,d}(\mathcal{I}_{N-2}^d) = \mathbb{E}[\mathbf{m}_{N-2}|\mathcal{I}_{N-2}^d]$. Similarly, the optimal decoder actions become $\mathbf{u}_k^* = \gamma_k^{*,d}(\mathcal{I}_k^d) = \mathbb{E}[\mathbf{m}_k|\mathcal{I}_k^d] = \mathbb{E}[\mathbf{m}_k|\mathbf{x}_{[0,k]}]$. Then, due to the Stackelberg assumption, the total encoder cost becomes $J^e(\gamma_{[0,N-1]}^e, \gamma_{[0,N-1]}^{*,d}) = \mathbb{E}\left[\sum_{k=0}^{N-1} \|\mathbf{m}_k - \mathbf{u}_k\|^2\right] + N\|\mathbf{b}\|^2$ by the smoothing property of the expectation. Thus, as in the static game setup (Sarıtaş et al., 2017b, Theorem 3.3), the goals of the players become essentially the same, and the result follows. □

3. Multi-stage quadratic Gaussian signaling games

The multi-stage signaling game setup is similar to the multi-stage cheap talk setup except that there exists an additive Gaussian noise channel between the encoder and the decoder at each stage, and the encoder has a *soft* power constraint. For the purpose of illustration, the system model of the 2-stage signaling game is depicted in Fig. 1(b). Here, source is assumed to be an n-dimensional Markovian source with initial Gaussian distribution; i.e., $\mathbf{M}_0 \sim \mathcal{N}(0, \Sigma_{\mathbf{M}_0})$ and $\mathbf{M}_{k+1} = G\mathbf{M}_k + \mathbf{V}_k$ where G is an $n \times n$ matrix (g denotes the scalar case equivalent of G) and $\mathbf{V}_k \sim \mathcal{N}(0, \Sigma_{\mathbf{V}_k})$ is an i.i.d. Gaussian noise sequence. The channels between the encoder and the decoder are assumed to be i.i.d. additive Gaussian channels; i.e., $\mathbf{W}_k \sim \mathcal{N}(0, \Sigma_{\mathbf{W}_k})$, and $\mathbf{W}_k$ and $\mathbf{V}_l$ are independent. At the kth stage of the N-stage game, the encoder knows the values of $\mathcal{I}_k^e = \{\mathbf{m}_{[0,k]}, \mathbf{y}_{[0,k-1]}\}$ (a noiseless feedback channel is assumed) and the decoder knows the values of $\mathcal{I}_k^d = \{\mathbf{y}_{[0,k]}\}$ with $\mathbf{y}_k = \mathbf{x}_k + \mathbf{w}_k$. Thus, under the policies considered, $\mathbf{x}_k = \gamma_k^e(\mathcal{I}_k^e)$ and $\mathbf{u}_k = \gamma_k^d(\mathcal{I}_k^d)$. The encoder's goal is to minimize (3) with $c_k^e(\mathbf{m}_k, \mathbf{x}_k, \mathbf{u}_k) = \|\mathbf{m}_k - \mathbf{u}_k - \mathbf{b}\|^2 + \lambda\|\mathbf{x}_k\|^2$, whereas, the decoder's goal is to minimize (4) with $c_k^d(\mathbf{m}_k, \mathbf{u}_k) = \|\mathbf{m}_k - \mathbf{u}_k\|^2$, by finding the optimal policy sequences $\gamma_{[0,N-1]}^{*,e}$ and $\gamma_{[0,N-1]}^{*,d}$, respectively, where the lengths of the vectors are defined in L_2 norm and $\mathbf{b}$ is the bias vector. Note that a power constraint with an associated multiplier λ 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.

3.1. Nash equilibrium analysis of multi-stage quadratic Gaussian signaling games[5]

In multi-stage quadratic Gaussian signaling games, affine policies constitute an invariant subspace under best response maps for Nash equilibria for both scalar ($n = 1$) and vector ($n > 1$) Gauss–Markov sources.

Theorem 11.

(i) *If the encoder uses affine policies at all stages, then the decoder will be affine at all stages.*

(ii) *If the decoder uses affine policies at all stages, then the encoder will be affine at all stages.*

Proof. (i) Let the encoder policies be $\mathbf{x}_k = \gamma_k^e(\mathbf{m}_{[0,k]}, \mathbf{y}_{[0,k-1]}) = \sum_{i=0}^{k} A_{k,i}\mathbf{m}_i + \sum_{i=0}^{k-1} B_{k,i}\mathbf{y}_i + \mathbf{C}_k$ where $A_{k,i}$ and $B_{k,i}$ are $n \times n$ matrices, and $\mathbf{C}_k$ is $n \times 1$ vector for $k \leq N - 1$ and $i \leq k$. Then, the optimal decoder actions are $\mathbf{u}_k^* = \mathbb{E}[\mathbf{m}_k|\mathcal{I}_k^d] = \mathbb{E}[\mathbf{m}_k|\mathbf{y}_{[0,k]}]$ for $k \leq N - 1$. Notice that $\mathbf{y}_{[0,k]}$ is multivariate Gaussian since $\mathbf{y}_k = \mathbf{x}_k + \mathbf{w}_k$. This proves that $\gamma_k^{*,d}(\mathcal{I}_k^d)$ is an affine function of $\mathbf{y}_{[0,k]}$ due to the joint Gaussianity of $\mathbf{m}_k$ and $\mathbf{y}_{[0,k]}$.

(ii) For a fixed affine decoder, the optimal encoder can be computed algebraically (for a detailed derivation, see Sarıtaş (2018, Theorem 3.5.1)). □

While it provides a structural result on the plausibility of affine equilibria, Theorem 11 does not lead to a conclusion about the existence of an informative equilibrium. It may be tempting to apply fixed point theorems (such as Brouwer's fixed point theorem Başar & Olsder, 1999) to establish the existence of informative equilibria; however, that there always exists a non-informative equilibrium for the cheap talk game also applies to the signaling game (Sarıtaş et al., 2017b). Later on, we will make information theoretic arguments (in Theorem 13) for the existence of informative equilibria for the Stackelberg setup, but this is not feasible for the Nash setup. However, the informativeness analysis of the 2-stage signaling-game can be accomplished by analyzing the fixed points of the invariant set of affine policies in Theorem 11 as follows (for a proof, due to space constraints, see Sarıtaş (2018, Theorem 3.4.2)):

Theorem 12. *For the 2-stage signaling game setup under affine encoder and decoder assumptions,*

(i) *If $\lambda > \max\left\{\frac{(g^2+1)\sigma_{M_0}^2}{\sigma_{W_0}^2}, \frac{\sigma_{M_1}^2}{\sigma_{W_1}^2}\right\}$, then there does not exist an informative affine equilibrium.*

(ii) *If $\frac{\sigma_{M_1}^2}{\sigma_{W_1}^2} < \lambda \leq \frac{(g^2+1)\sigma_{M_0}^2}{\sigma_{W_0}^2}$, then the second stage message m_1 is not used in the game.*

(iii) *If $\frac{(g^2+1)\sigma_{M_0}^2}{\sigma_{W_0}^2} < \lambda \leq \frac{\sigma_{M_1}^2}{\sigma_{W_1}^2}$, the equilibrium is informative if and only if $\sigma_{M_1}^2 \geq 4b^2$ and $\max\left\{\frac{\sigma_{M_1}^2 - 2b^2 - \sqrt{\sigma_{M_1}^2}\sqrt{\sigma_{M_1}^2 - 4b^2}}{2\sigma_{W_1}^2},\right.$*

$\left.(g^2 + 1)\frac{\sigma_{M_0}^2}{\sigma_{W_0}^2}\right\} < \lambda < \frac{\sigma_{M_1}^2 - 2b^2 + \sqrt{\sigma_{M_1}^2}\sqrt{\sigma_{M_1}^2 - 4b^2}}{2\sigma_{W_1}^2}.$

The above analysis can be carried over to the N-stage signaling game; however, this would involve $(3N^2 + 5N)/2$ equations and as many unknowns.

3.2. Stackelberg equilibrium analysis of multi-stage quadratic Gaussian signaling games

Here, the signaling game is analyzed under the Stackelberg concept.

3.2.1. Multi-stage Stackelberg equilibria for scalar Gauss–Markov sources

The conditions for informative Stackelberg equilibria with scalar sources are characterized below. The proof is in Appendix A.

Theorem 13. *If $\lambda \geq \max_{k \leq N-1} \frac{\sigma_{M_k}^2}{\sigma_{W_k}^2} \sum_{i=0}^{N-k-1} g^{2i}$, there does not exist an informative (affine or non-linear) equilibrium in the N-stage scalar signaling game under the Stackelberg assumption; i.e., the only equilibrium is the non-informative one. Otherwise, an equilibrium has to be always linear.*

Now consider the multi-stage Stackelberg signaling game with a discounted infinite horizon and a discount factor $\beta \in (0, 1)$; i.e., $J^e(\gamma^e, \gamma^d) = \mathbb{E}\left[\sum_{i=0}^{\infty} \beta^i \left((m_i - u_i - b)^2 + \lambda x_i^2\right)\right]$ and $J^d(\gamma^e, \gamma^d) = \mathbb{E}\left[\sum_{i=0}^{\infty} \beta^i (m_i - u_i)^2\right]$. The proof is in Appendix B.

Corollary 14. *If* $\lambda \geq \max_{k=0,1,\dots} \frac{\sigma_{M_k}^2}{\sigma_{W_k}^2} \frac{1}{1-\beta g^2}$ *where* $\beta g^2 < 1$, *there does not exist an informative (affine or non-linear) equilibrium in the infinite horizon discounted multi-stage Stackelberg signaling game for scalar Gauss–Markov sources.*

3.2.2. Multi-stage Stackelberg equilibria for vector Gauss–Markov sources

Linear policies are optimal for scalar sources as shown in Section 3.2.1. Before analyzing the multi-dimensional setup, it will be appropriate to review the optimality of linear policies in Gaussian setups for the classical communication theoretic setup when the bias term is absent: Optimality of linear coding policies for scalar Gaussian source–channel pairs with noiseless feedback has been known since 1960s, see e.g. Goblick (1965). Optimal linear encoders for single-stage setups have been studied in Başar (1978) and Başar (1980). When the source and the channel are multi-dimensional, linear policies may not be optimal; see Pilc (1969), Yüksel and Başar (2013, Chapter 11) and Zaidi, Oechtering, Yüksel, and Skoglund (2013) for a detailed discussion and literature review. It is evident from Theorem 11 that when the encoder is linear, the optimal decoder is linear. In this case, a relevant problem is to find the optimal Stackelberg policy among the linear or affine class. In the following, a dynamic programming approach is adapted to find such Stackelberg equilibria. Building on the optimality of linear innovation encoders, we restrict the analysis to such encoders; i.e., we consider a sub-optimal scenario. Our analysis builds on and generalizes the arguments in Başar and Bansal (1994) and Zaidi et al. (2013, Theorem 3). The proof is in Appendix C.

Theorem 15. *Suppose that G, $\Sigma_{\mathbf{M}_0}$ and $\Sigma_{\mathbf{v}_k}$ are diagonal. Suppose further that the innovation is given by $\widetilde{\mathbf{m}}_k \triangleq \mathbf{m}_k - \mathbb{E}[\mathbf{m}_k|\mathbf{y}_{[0,k-1]}]$ with $\widetilde{\mathbf{m}}_0 = \mathbf{m}_0$, and that the encoder linearly encodes the innovation. Then, an optimal such linear policy can be computed through dynamic programming with value functions $V_k\left(\Sigma_{\widetilde{\mathbf{m}}_k}\right) \triangleq \mathrm{tr}\left(K_k \Sigma_{\widetilde{\mathbf{m}}_k} + L_k\right)$ that satisfy the terminal condition $V_N\left(\Sigma_{\widetilde{\mathbf{m}}_N}\right) = 0$ with diagonal K_k matrices for $k \leq N - 1$.*

4. Concluding remarks

In this paper, we studied Nash and Stackelberg equilibria for multi-stage quadratic cheap talk and signaling games. We established qualitative (e.g. on full revelation, quantization nature, linearity, informativeness and non-informativeness) and quantitative properties (on linearity or explicit computation) of Nash and Stackelberg equilibria under misaligned objectives.

Appendix A. Proof of Theorem 13

Similar to that in Theorem 10, the optimal decoder actions are $u_k^* = \gamma_k^{*,d}(\mathcal{I}_k^d) = \mathbb{E}[m_k|\mathcal{I}_k^d] = \mathbb{E}[m_k|y_{[0,k]}]$, and the total encoder cost becomes $J^e(\gamma_{[0,N-1]}^e, \gamma_{[0,N-1]}^d) = \mathbb{E}\left[\sum_{k=0}^{N-1} \mathbb{E}[(m_k - \mathbb{E}[m_k|\mathcal{I}_k^d])^2 + b^2 + \lambda x_k^2|\mathcal{I}_k^d]\right]$. This is an instance of problems studied in Bansal and Başar (1989), and can be reduced to a team problem where both the players are minimizing the same cost. The linearity of the optimal policies can be deduced from Bansal and Başar (1989). Here, we adapt the proof in Bansal and Başar (1989) to our setup. From the chain rule, $I(m_k; y_{[0,k]}) = I(m_k; y_{[0,k-1]}) + I(m_k; y_k|y_{[0,k-1]})$. By following similar arguments to those in Bansal and Başar (1989) and Yüksel and Başar (2013, Theorem 11.3.1),

$$I(m_k; y_k|y_{[0,k-1]}) = h(y_k|y_{[0,k-1]}) - h(y_k|m_k, y_{[0,k-1]})$$

$$\leq h(y_k) - h(y_k|\gamma_k^e(m_k, y_{[0,k-1]}))$$
$$= I\left(\gamma_k^e(m_k, y_{[0,k-1]}); y_k\right) = I(x_k; y_k) \leq \sup I(x_k; y_k)$$
$$= \frac{1}{2}\log_2\left(1 + \frac{P_k}{\sigma_{W_k}^2}\right) \triangleq \widehat{C}_k \text{ where } P_k = \mathbb{E}[x_k^2].$$

It can be seen that $m_k - \mathbb{E}[m_k|m_{k-1}]$ is orthogonal to the random variables $m_{k-1}, y_{[0,k-1]}$ where $y_{[0,k-1]}$ is included due to the Markov chain $m_k \leftrightarrow m_{k-1} \leftrightarrow (y_{[0,k-1]})$. By using this orthogonality, it follows that

$$
\begin{aligned}
\mathbb{E}[(m_k - \mathbb{E}[m_k|y_{[0,k-1]}])^2] &= \mathbb{E}[(m_k - \mathbb{E}[m_k|m_{k-1}])^2] \\
&\quad + \mathbb{E}[(\mathbb{E}[m_k|m_{k-1}] - \mathbb{E}[m_k|y_{[0,k-1]}])^2] \\
&\overset{(a)}{=} \mathbb{E}[(m_k - \mathbb{E}[m_k|m_{k-1}])^2] + \mathbb{E}\Big[\big(\mathbb{E}[m_k|m_{k-1}] \\
&\qquad - \mathbb{E}[\mathbb{E}[m_k|m_{k-1}]|y_{[0,k-1]}]\big)^2\Big] \\
&\overset{(b)}{=} \sigma_{V_{k-1}}^2 + g^2 \mathbb{E}[(m_{k-1} - \mathbb{E}[m_{k-1}|y_{[0,k-1]}])^2] \\
&\overset{(c)}{\geq} \sigma_{V_{k-1}}^2 + g^2 \sigma_{M_{k-1}}^2 2^{-2C_{k-1}}, \quad (A.1)
\end{aligned}
$$

where $C_k \triangleq \sup I(m_k; y_{[0,k]})$. Here, (a) holds due to the iterated expectation rule and the Markov chain property, (b) holds since $\mathbb{E}[m_k|m_{k-1}] = \mathbb{E}[gm_{k-1} + v_{k-1}|m_{k-1}] = gm_{k-1}$, and (c) holds due to Yüksel and Başar (2013, Lemma 11.3.1). From Yüksel and Başar (2013, Lemma 11.3.2), $I(m_k; y_{[0,k-1]})$ is maximized with linear policies, and the lower bound of (A.1), $\mathbb{E}[(m_k - \mathbb{E}[m_k|y_{[0,k-1]}])^2] \geq \sigma_{V_{k-1}}^2 + g^2 \sigma_{M_{k-1}}^2 2^{-2C_{k-1}} \triangleq \sigma_{M_k}^2 2^{-2\widetilde{C}_k}$, is achievable through linear policies where $\sup I(m_k; y_{[0,k-1]}) \triangleq \widetilde{C}_k = \frac{1}{2}\log_2\left(\frac{\sigma_{M_k}^2}{\sigma_{V_{k-1}}^2 + g^2 \sigma_{M_{k-1}}^2 2^{-2C_{k-1}}}\right)$. Thus, we have the following recursion on upper bounds on mutual information for the N-stage signaling game:

$$
\begin{aligned}
C_k &= \sup I(m_k; y_{[0,k]}) = \widetilde{C}_k + \widehat{C}_k \\
&= \frac{1}{2}\log_2\left(\frac{\sigma_{M_k}^2}{\sigma_{V_{k-1}}^2 + g^2 \sigma_{M_{k-1}}^2 2^{-2C_{k-1}}}\right) + \frac{1}{2}\log_2\left(1 + \frac{P_k}{\sigma_{W_k}^2}\right)
\end{aligned}
$$

for $k \leq N - 1$ with $C_0 = \frac{1}{2}\log_2\left(1 + \frac{P_0}{\sigma_{W_0}^2}\right)$. Let the lower bound on $\mathbb{E}\left[(m_k - \mathbb{E}[m_k|y_{[0,k]}])^2\right]$ be Δ_k; i.e., $\mathbb{E}\left[(m_k - \mathbb{E}[m_k|y_{[0,k]}])^2\right] \geq \sigma_{M_k}^2 2^{-2C_k} \triangleq \Delta_k$. Then the following recursion holds for the N-stage signaling game:

$$\Delta_k = \frac{\sigma_{V_{k-1}}^2 + g^2 \Delta_{k-1}}{1 + \frac{P_k}{\sigma_{W_k}^2}} \text{ for } k = 1, 2, \dots, N - 1$$

with $\Delta_0 = \frac{\sigma_{M_0}^2}{1 + \frac{P_0}{\sigma_{W_0}^2}}$. Since $\Delta_k = \sigma_{M_k}^2 2^{-2C_k}$ by definition, $\Delta_k \leq \sigma_{M_k}^2$ for $k \leq N - 1$. In an equilibrium, since the decoder always chooses $u_k = \mathbb{E}[m_k|y_{[0,k]}]$ for $k \leq N - 1$, the total encoder cost for the first stage can be lower bounded by $J_0^{e,lower} = \sum_{i=0}^{N-1}\left(\Delta_i + \lambda P_i + b^2\right)$. Now observe the following:

$$
\frac{\partial \Delta_l}{\partial P_k} = \begin{cases}
0 & \text{if } l < k \\
g^2\left(1 + \frac{P_l}{\sigma_{W_l}^2}\right)^{-1}\frac{\partial \Delta_{l-1}}{\partial P_k} - \frac{1}{\sigma_{W_l}^2}\frac{\partial P_l}{\partial P_k} \\
\quad \times \left(\sigma_{V_{l-1}}^2 + g^2 \Delta_{l-1}\right)\left(1 + \frac{P_l}{\sigma_{W_l}^2}\right)^{-2} & \text{if } l \geq k
\end{cases},
$$

where $\frac{\partial P_l}{\partial P_k} = 0$ for $l < k$ due to the information structure of the encoder. Then, $\frac{\partial J_0^{e,lower}}{\partial P_{N-1}} \geq \lambda - \frac{\sigma_{M_{N-1}}^2}{\sigma_{W_{N-1}}^2}$ can be obtained. If $\lambda > \frac{\sigma_{M_{N-1}}^2}{\sigma_{W_{N-1}}^2}$, then $\frac{\partial J_0^{e,lower}}{\partial P_{N-1}} > 0$, which implies that $J_0^{e,lower}$ is an increasing function of P_{N-1}. For this case, in order to minimize $J_0^{e,lower}$, P_{N-1} must be chosen as 0; i.e., $P_{N-1}^* = 0$. Then, for $\lambda > \frac{\sigma_{M_{N-1}}^2}{\sigma_{W_{N-1}}^2}$,

$$\frac{\partial J_0^{e,lower}}{\partial P_{N-2}} = \lambda \left(1 + \frac{\partial P_{N-1}}{\partial P_{N-2}}\right) + \sum_{i=N-2}^{N-1} \frac{\partial \Delta_i}{\partial P_{N-2}}$$

$$= \lambda \left(1 + \frac{\partial P_{N-1}}{\partial P_{N-2}}\right) + \left(g^2 \left(1 + \frac{P_{N-1}}{\sigma_{W_{N-1}}^2}\right)^{-1} + 1\right) \frac{\partial \Delta_{N-2}}{\partial P_{N-2}}$$

$$- \left(\sigma_{V_{N-2}}^2 + g^2 \Delta_{N-2}\right) \left(1 + \frac{P_{N-1}}{\sigma_{W_{N-1}}^2}\right)^{-2} \frac{1}{\sigma_{W_{N-1}}^2} \frac{\partial P_{N-1}}{\partial P_{N-2}}$$

$$\overset{(a)}{=} \lambda + \frac{\partial \Delta_{N-2}}{\partial P_{N-2}} \left(g^2 + 1\right) \geq \lambda - \frac{\sigma_{M_{N-2}}^2}{\sigma_{W_{N-2}}^2} \left(g^2 + 1\right).$$

Here, (a) holds since $P_{N-1}^* = 0$ for $\lambda > \frac{\sigma_{M_{N-1}}^2}{\sigma_{W_{N-1}}^2}$. If $\lambda > \max\left\{\frac{\sigma_{M_{N-1}}^2}{\sigma_{W_{N-1}}^2}, \frac{\sigma_{M_{N-2}}^2}{\sigma_{W_{N-2}}^2} \left(g^2 + 1\right)\right\}$, then $\frac{\partial J_0^{e,lower}}{\partial P_{N-2}} > 0$, which implies that $J_0^{e,lower}$ is an increasing function of P_{N-2}. For this case, in order to minimize $J_0^{e,lower}$, P_{N-2} must be chosen as 0. By following the similar approach and assumptions on λ, since $P_{N-1}^* = P_{N-2}^* = \cdots = P_{k+1}^* = 0$, we have the following:

$$\frac{\partial J_0^{e,lower}}{\partial P_k} = \lambda + \sum_{i=k}^{N-1} \frac{\partial \Delta_i}{\partial P_k} = \lambda + \frac{\partial \Delta_k}{\partial P_k} \sum_{i=k}^{N-1} \prod_{j=k+1}^{i} g^2$$

$$= \lambda - \frac{\sigma_{V_{k-1}}^2 + g^2 \Delta_{k-1}}{\sigma_{W_k}^2} \left(1 + \frac{P_k}{\sigma_{W_k}^2}\right)^{-2} \sum_{i=k}^{N-1} \prod_{j=k+1}^{i} g^2$$

$$\geq \lambda - \frac{\sigma_{M_k}^2}{\sigma_{W_k}^2} \sum_{i=0}^{N-k-1} g^{2i},$$

where $\prod_{i=k}^{l} = 1$ if $k > l$. If $\lambda > \frac{\sigma_{M_k}^2}{\sigma_{W_k}^2} \sum_{i=0}^{N-k-1} g^{2i}$, then $\frac{\partial J_0^{e,lower}}{\partial P_k} > 0$, which implies that $J_0^{e,lower}$ is an increasing function of P_k. For this case, in order to minimize $J_0^{e,lower}$, P_k must be chosen as 0.

By combining all the results above, it can be deduced that if $\lambda > \max_{k \leq N-1} \frac{\sigma_{M_k}^2}{\sigma_{W_k}^2} \sum_{i=0}^{N-k-1} g^{2i}$, the lower bound $J_0^{e,lower}$ of the encoder costs J_0^e is minimized by choosing $P_0^* = P_1^* = \cdots = P_{N-1}^* = 0$; that is, the encoder does not signal any output. Hence, the encoder engages in a non-informative equilibrium and the minimum cost becomes $J_0^e = J_0^{e,lower} = \left(\sum_{i=0}^{N-1} \sigma_{M_i}^2\right) + Nb^2$. $\square$

Appendix B. Proof of Corollary 14

For the infinite horizon case, it can be observed

$$\inf_{\gamma^e} \limsup_{N \to \infty} J^e(\gamma_{[0,N-1]}^e, \gamma_{[0,N-1]}^d)$$

$$\geq \limsup_{N \to \infty} \inf_{\gamma_{[0,N-1]}^e} \sum_{i=0}^{N-1} \beta^i \left(\Delta_i + \lambda P_i + b^2\right).$$

Thus, $\limsup_{N \to \infty} \inf_{\gamma_{[0,N-1]}^e} \sum_{i=0}^{N-1} \beta^i \left(\Delta_i + \lambda P_i + b^2\right)$ is achieved at a non-informative equilibrium if $\lambda > \limsup_{N \to \infty} \max_{k \leq N-1}$

$\frac{\sigma_{M_k}^2}{\sigma_{W_k}^2} \sum_{i=0}^{N-k-1} \beta^i g^{2i} = \frac{\sigma_{M_k}^2}{\sigma_{W_k}^2} \frac{1}{1 - \beta g^2}$ for $\beta g^2 < 1$. Hence, if $\lambda \geq \max_{k=0,1,\ldots} \frac{\sigma_{M_k}^2}{\sigma_{W_k}^2} \frac{1}{1 - \beta g^2}$, then the lower bound $J_0^{e,lower}$ of the encoder costs J_0^e is minimized by choosing $P_0 = P_1 = \cdots = 0$, and the minimum cost becomes $J_0^e = J_0^{e,lower} = \sum_{i=0}^{\infty} \beta^i \left(\sigma_{M_i}^2 + b^2\right)$ at this non-informative equilibrium. $\square$

Appendix C. Proof of Theorem 15

We will follow an approach similar to that in Zaidi et al. (2013) which restricted the analysis to a team problem and a scalar channel; Zaidi et al. (2013) in turn build on Başar and Bansal (1994), which considers continuous time systems. Since the $(k+1)$st stage encoder policy only transmits the linearly encoded innovation by assumption, $\mathbf{x}_k = \gamma_k^e(\mathcal{I}_k^e) = A_k \widetilde{\mathbf{m}}_k$ where A_k is an $n \times n$ matrix for $k \leq N - 1$. Then the decoder receives $\mathbf{y}_k = \mathbf{x}_k + \mathbf{w}_k = A_k \widetilde{\mathbf{m}}_k + \mathbf{w}_k$ and applies the action $\mathbf{u}_k = \gamma_k^d(\mathcal{I}_k^d) = \mathbb{E}[\mathbf{m}_k | \mathbf{y}_{[0,k]}]$ to minimize his stage-wise cost $\|\mathbf{e}_k\|^2 \triangleq \mathbb{E}[\|\mathbf{m}_k - \mathbf{u}_k\|^2] = \mathbb{E}[(\mathbf{m}_k - \mathbf{u}_k)^T(\mathbf{m}_k - \mathbf{u}_k)] = \text{tr}\left(\Sigma_{\mathbf{e}_k}\right)$ for $k \leq N - 1$ where $\Sigma_{\mathbf{R}}$ stands for the covariance matrix of the random variable $\mathbf{R}$; i.e., $\Sigma_{\mathbf{R}} \triangleq \mathbb{E}[(\mathbf{R} - \mathbb{E}[\mathbf{R}])(\mathbf{R} - \mathbb{E}[\mathbf{R}])^T]$. Due to the orthogonality of $\widetilde{\mathbf{m}}_k$ and $\mathbf{y}_{[0,k-1]}$, and the iterated expectations rule, $\mathbf{u}_k = \mathbb{E}[\mathbf{m}_k | \mathbf{y}_{[0,k]}] = \mathbb{E}\left[\widetilde{\mathbf{m}}_k + \mathbb{E}[\mathbf{m}_k | \mathbf{y}_{[0,k-1]}] | \mathbf{y}_{[0,k]}\right] = \mathbb{E}[\widetilde{\mathbf{m}}_k | \mathbf{y}_k] + \mathbb{E}[\mathbf{m}_k | \mathbf{y}_{[0,k-1]}]$, and it follows that $\mathbf{e}_k = \mathbf{m}_k - \mathbf{u}_k = \mathbf{m}_k - \mathbb{E}[\widetilde{\mathbf{m}}_k | \mathbf{y}_k] - \mathbb{E}[\mathbf{m}_k | \mathbf{y}_{[0,k-1]}] = \widetilde{\mathbf{m}}_k - \mathbb{E}[\widetilde{\mathbf{m}}_k | \mathbf{y}_k]$. Since $\mathbb{E}[\widetilde{\mathbf{m}}_k | \mathbf{y}_k] = \Sigma_{\widetilde{\mathbf{m}}_k} A_k^T \left(\Sigma_{\mathbf{y}_k}\right)^{-1} \mathbf{y}_k$, the stage-wise cost of the decoder becomes the trace of the following:

$$\Sigma_{\mathbf{e}_k} = \Sigma_{\widetilde{\mathbf{m}}_k} - \Sigma_{\widetilde{\mathbf{m}}_k} A_k^T \left(A_k \Sigma_{\widetilde{\mathbf{m}}_k} A_k^T + \Sigma_{\mathbf{w}_k}\right)^{-1} A_k \Sigma_{\widetilde{\mathbf{m}}_k}$$
$$\overset{(a)}{=} \Sigma_{\widetilde{\mathbf{m}}_k}^{1/2} \left(I + H_k^T H_k\right)^{-1} \Sigma_{\widetilde{\mathbf{m}}_k}^{1/2}, \tag{C.1}$$

where (a) follows by utilizing the matrix inversion lemma, $(I + UWV)^{-1} = I - U(W^{-1} + VU)^{-1}V$, where $U = H_k^T$, $W = I$, $V = H_k$, and $H_k \triangleq \Sigma_{\mathbf{w}_k}^{-1/2} A_k \Sigma_{\widetilde{\mathbf{m}}_k}^{1/2}$. Since $\mathbb{E}[\mathbf{m}_k | \mathbf{y}_{[0,k-1]}]$ and $\widetilde{\mathbf{m}}_k$ are orthogonal, $\mathbb{E}[\mathbf{m}_k | \mathbf{y}_k] = \Sigma_{\widetilde{\mathbf{m}}_k} A_k^T \left(\Sigma_{\mathbf{y}_k}\right)^{-1} \mathbf{y}_k$. Then the innovation and its covariance matrix can be expressed recursively as follows:

$$\widetilde{\mathbf{m}}_{k+1} = G\mathbf{m}_k + \mathbf{v}_k - \mathbb{E}[\mathbf{m}_{k+1} | \mathbf{y}_{[0,k-1]}] - \mathbb{E}[\mathbf{m}_{k+1} | \mathbf{y}_k]$$
$$= G\widetilde{\mathbf{m}}_k + \mathbf{v}_k - G\Sigma_{\widetilde{\mathbf{m}}_k} A_k^T \left(\Sigma_{\mathbf{y}_k}\right)^{-1} \mathbf{y}_k,$$

$$\Sigma_{\widetilde{\mathbf{m}}_{k+1}} = G\Sigma_{\widetilde{\mathbf{m}}_k}^{1/2} \left(I + H_k^T H_k\right)^{-1} \Sigma_{\widetilde{\mathbf{m}}_k}^{1/2} G^T + \Sigma_{\mathbf{v}_k}. \tag{C.2}$$

Further, the stage-wise cost of the encoder is

$$\mathbb{E}\left[\|\mathbf{m}_k - \mathbf{u}_k - \mathbf{b}\|^2 + \lambda \|\mathbf{x}_k\|^2\right] = \text{tr}\left(\Sigma_{\mathbf{e}_k}\right) + \text{tr}\left(\lambda \Sigma_{\mathbf{x}_k}\right) + \|\mathbf{b}\|^2$$
$$= \text{tr}\left(\Sigma_{\widetilde{\mathbf{m}}_k} \left(I + H_k^T H_k\right)^{-1}\right) + \text{tr}\left(\lambda H_k^T \Sigma_{\mathbf{w}_k} H_k\right) + \|\mathbf{b}\|^2. \tag{C.3}$$

Let the value functions be $V_k\left(\Sigma_{\widetilde{\mathbf{m}}_k}\right) = \text{tr}\left(K_k \Sigma_{\widetilde{\mathbf{m}}_k} + L_k\right)$ with K_k being diagonal. In the following we show that there exist such V_k that satisfy Bellman's principle of optimality (Hernandez-Lerma & Lasserre, 1996, Theorem 3.2.1). Here, $V_k\left(\Sigma_{\widetilde{\mathbf{m}}_k}\right) \triangleq \min_{H_k} \left(\mathcal{C}_k\left(\Sigma_{\widetilde{\mathbf{m}}_k}, H_k\right) + V_{k+1}\left(\Sigma_{\widetilde{\mathbf{m}}_{k+1}}\right)\right)$, and $\mathcal{C}_k\left(\Sigma_{\widetilde{\mathbf{m}}_k}, H_k\right) \triangleq \text{tr}\left(\Sigma_{\mathbf{e}_k}\right) + \text{tr}\left(\lambda \Sigma_{\mathbf{x}_k}\right) + \|\mathbf{b}\|^2$ is the stage-wise cost of the kth stage encoder. Then, since K_{k+1} and L_{k+1} do not depend on H_k,

$$V_k\left(\Sigma_{\widetilde{\mathbf{m}}_k}\right) = \min_{H_k} \left(\mathcal{C}_k\left(\Sigma_{\widetilde{\mathbf{m}}_k}, H_k\right) + V_{k+1}\left(\Sigma_{\widetilde{\mathbf{m}}_{k+1}}\right)\right)$$

$$\overset{(a)}{=} \text{tr}\left(K_{k+1} \Sigma_{\mathbf{v}_k} + L_{k+1}\right) + \|\mathbf{b}\|^2 + \min_{H_k} \Big(\underbrace{\text{tr}\left(\lambda H_k^T \Sigma_{\mathbf{w}_k} H_k\right)}_{\triangleq \mathfrak{C}}$$

$$
+ \; \mathrm{tr}\Big(\underbrace{\Sigma_{\widetilde{\mathbf{m}}_k}^{1/2} \left(G^T K_{k+1} G + I \right) \Sigma_{\widetilde{\mathbf{m}}_k}^{1/2} \left(I + H_k^T H_k \right)^{-1} }_{\triangleq \mathfrak{P}} \Big) \Big) , \tag{C.4}
$$

where (a) follows by substituting $\mathcal{C}_k \left(\Sigma_{\widetilde{\mathbf{m}}_k}, H_k \right)$ using (C.3) and employing (C.2). The equivalent problem of the minimization of $\mathfrak{P}$ over H_k under the constraint $\mathfrak{C} = \mu_k$ is considered in Başar (1980), and for every $\mu_k = \mathfrak{C}$, the optimal H_k is found as $H_k^* = \Pi_k \zeta_k P_k^T$, where Π_k is a unitary matrix such that $\Pi_k^T \left(\lambda \Sigma_{\mathbf{w}_k} \right) \Pi_k = \mathrm{diag}\left(\tau_{k_1}, \tau_{k_2}, \ldots, \tau_{k_n} \right) \triangleq \widetilde{\Pi}_k$, P_k is a unitary matrix such that $P_k^T \left(\Sigma_{\widetilde{\mathbf{m}}_k}^{1/2} \left(G^T K_{k+1} G + I \right) \Sigma_{\widetilde{\mathbf{m}}_k}^{1/2} \right) P_k = \mathrm{diag}\left(\nu_{k_1}, \nu_{k_2}, \ldots, \nu_{k_n} \right)$, and ζ_k is another diagonal matrix. Then the recursion of the innovation's covariance matrix (C.2) can be expressed as

$$
\Sigma_{\widetilde{\mathbf{m}}_{k+1}} = G \Sigma_{\widetilde{\mathbf{m}}_k}^{1/2} \left(I + P_k \zeta_k^T \zeta_k P_k^T \right)^{-1} \Sigma_{\widetilde{\mathbf{m}}_k}^{1/2} G^T + \Sigma_{\mathbf{v}_k} . \tag{C.5}
$$

Then, by utilizing $H_k^* = \Pi_k \zeta_k P_k^T$, $\Pi_k^T \Pi_k = I$, $\widetilde{\Pi}_k = \Pi_k^T \left(\lambda \Sigma_{\mathbf{w}_k} \right) \Pi_k$, and $P_k^T P_k = I$ in (C.4),

$$
V_k \left(\Sigma_{\widetilde{\mathbf{m}}_k} \right) \overset{(a)}{=} \mathrm{tr}\left(K_{k+1} \Sigma_{\mathbf{v}_k} + L_{k+1} + \mathbf{b}\mathbf{b}^T \right) + \mathrm{tr}\left(\zeta_k^T \widetilde{\Pi}_k \zeta_k \right)
$$
$$
+ \; \mathrm{tr}\left(\left(G^T K_{k+1} G + I \right) \left(I - \zeta_k^T \left(I + \zeta_k \zeta_k^T \right)^{-1} \zeta_k \right) \Sigma_{\widetilde{\mathbf{m}}_k} \right) , \tag{C.6}
$$

where (a) follows from the matrix inversion lemma by choosing $U = P_k \zeta_k^T$, $W = I$, and $V = \zeta_k P_k^T$ in $(I + UWV)^{-1} = I - U(W^{-1} + VU)^{-1}V$, and the diagonality of $\Sigma_{\widetilde{\mathbf{m}}_k}$, P_k and ζ_k: Since G, $\Sigma_{\widetilde{\mathbf{m}}_0}$, K_k and $\Sigma_{\mathbf{v}_k}$ are diagonal for $k \leq N - 1$, it is always possible to find a unitary diagonal P_0 such that $P_0^T \left(\Sigma_{\widetilde{\mathbf{m}}_0}^{1/2} \left(G^T K_1 G + I \right) \Sigma_{\widetilde{\mathbf{m}}_0}^{1/2} \right) P_0 = \mathrm{diag}\left(\nu_{0_1}, \nu_{0_2}, \ldots, \nu_{0_n} \right)$, which makes $\Sigma_{\widetilde{\mathbf{m}}_1}$ diagonal by (C.5). Similarly, $\Sigma_{\widetilde{\mathbf{m}}_k}$ and P_k are diagonal for $k \leq N - 1$. In order to satisfy (C.6), since $V_N \left(\Sigma_{\widetilde{\mathbf{m}}_N} \right) = 0$, we choose $K_N = L_N = 0$. Then, for $k \leq N - 1$, $\{K_{k+1}, L_{k+1}\}$ is chosen according to

$$
K_k = \left(G^T K_{k+1} G + I \right) \left(I - \zeta_k^T \left(I + \zeta_k \zeta_k^T \right)^{-1} \zeta_k \right) ,
$$
$$
L_k = K_{k+1} \Sigma_{\mathbf{v}_k} + L_{k+1} + \zeta_k^T \widetilde{\Pi}_k \zeta_k + \mathbf{b}\mathbf{b}^T . \tag{C.7}
$$

Now we verify that the diagonal K_k matrices satisfy the dynamic programming recursion.

When the channel is scalar, for $k \leq N - 1$,

$$
K_k = \left(G^T K_{k+1} G + I \right) \times \; \mathrm{diag}\left(\frac{\lambda \sigma_{W_k}^2}{1 + \lambda \sigma_{W_k}^2}, 1, 1, \ldots, 1 \right) ,
$$
$$
L_k = K_{k+1} \Sigma_{\mathbf{v}_k} + L_{k+1} + \; \mathrm{diag}\left(1, 0, 0, \ldots, 0 \right) + \mathbf{b}\mathbf{b}^T .
$$

The optimal linear encoder policy is $A_k^* = \Sigma_{\mathbf{w}_k}^{1/2} \zeta_k P_k^T \Sigma_{\widetilde{\mathbf{m}}_k}^{-1/2}$ since $\Pi_k = 1$ and $\zeta_k = \left[\frac{1}{\sqrt{\lambda \sigma_{W_k}^2}}, 0, \ldots, 0 \right]$. $\square$

References

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

Aumann, R. J., & Hart, S. (2003). Long cheap talk. *Econometrica*, *71*(6), 1619–1660.

Başar, T. Ü. (1978). *Performance bounds and optimal linear coding for multichannel communication systems* (Ph.D. Dissertation), Boğaziçi University.

Başar, T. (1980). A trace minimization problem with applications in joint estimation and control under nonclassical information. *Journal of Optimization Theory and Applications*, *31*, 343–359.

Başar, T., & Bansal, R. (1994). Optimum design of measurement channels and control policies for linear-quadratic stochastic systems. *European Journal of Operational Research*, *73*(2), 226–236.

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

Bansal, R., & Başar, T. (1989). Simultaneous design of measurement and control strategies for stochastic systems with feedback. *Automatica*, *45*(5), 679–694.

Blackwell, D. (1951). The comparison of experiments. In *Proceedings of the second Berkeley symposium on mathematical statistics and probability* (pp. 93–102).

Borkar, V. S. (1993). White-noise representations in stochastic realization theory. *SIAM Journal on Control and Optimization (SICON)*, *31*, 1093–1102.

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

Ely, J., Frankel, A., & Kamenica, E. (2015). Suspense and surprise. *Journal of Political Economy*, *123*(1), 215–260.

Farokhi, F., Teixeira, A. M. H., & Langbort, C. (2017). Estimation with strategic sensors. *IEEE Transactions on Automatic Control*, *62*(2), 724–739.

Fleischer, P. E. (1964). Sufficient conditions for achieving minimum distortion in a quantizer. *IEEE International Convention Record, Part I*, *12*, 104–111.

Fudenberg, D., & Tirole, J. (1991). Perfect Bayesian equilibrium and sequential equilibrium. *Journal of Economic Theory*, *53*(2), 236–260.

Gihman, I. I., & Skorohod, A. V. (1979). *Controlled stochastic processes*. New York, NY: Springer-Verlag New York.

Goblick, T. (1965). Theoretical limitations on the transmission of data from analog sources. *IEEE Transactions on Information Theory*, *11*, 558–567.

Golosov, M., Skreta, V., Tsyvinski, A., & Wilson, A. (2014). Dynamic strategic information transmission. *Journal of Economic Theory*, *151*, 304–341.

Hernandez-Lerma, O., & Lasserre, J. (1996). *Discrete-time Markov control processes*. Springer.

Hirshleifer, J. (1971). The private and social value of information and the reward to inventive activity. *The American Economic Review*, 561–574.

Kamenica, E., & Gentzkow, M. (2011). Bayesian persuasion. *American Economic Review*, *101*(6), 2590–2615.

Kieffer, J. C. (1993). A survey of the theory of source coding with a fidelity criterion. *IEEE Transactions on Information Theory*, *39*(5), 1473–1490.

Korzhyk, D., Yin, Z., Kiekintveld, C., Conitzer, V., & Tambe, M. (2011). Stackelberg vs. Nash in security games: An extended investigation of interchangeability, equivalence, and uniqueness. *Journal of Artificial Intelligence Research*, *41*(2), 297–327.

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

Pilc, R. (1969). The optimum linear modulator for a Gaussian source used with a Gaussian channel. *IEEE Transactions on Automatic Cotrol*, *48*, 3075–3089.

Sarıtaş, S. (2018). *Signaling games in networked systems* (Ph.D. Dissertation), Bilkent University.

Sarıtaş, S., Furrer, P., Gezici, S., Linder, T., & Yüksel, S. (2019). On the number of bins in equilibria for signaling games. In *IEEE international symposium on information theory (ISIT)*, Paris, France (pp. 972–976).

Sarıtaş, S., Yüksel, S., & Gezici, S. (2016). Dynamic signaling games under Nash and Stackelberg equilibria. In *IEEE international symposium on information theory (ISIT)*, Barcelona, Spain (pp. 1631–1635).

Sarıtaş, S., Yüksel, S., & Gezici, S. (2017a). Nash and Stackelberg equilibria for dynamic cheap talk and signaling games. In *American control conference (ACC)* (pp. 3644–3649).

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

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

Shames, I., Teixeira, A. M. H., Sandberg, H., & Johansson, K. H. (2012). Agents misbehaving in a network: a vice or a virtue?. *IEEE Network*, *26*(3), 35–40.

Treust, M. L., & Tomala, T. (2018). Information-theoretic limits of strategic communication. arXiv preprint arXiv:1807.05147.

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

Vasal, D., & Anastasopoulos, A. (2016). Signaling equilibria for dynamic LQG games with asymmetric information. In *55th IEEE Conference on Decision and Control (CDC)*, Las Vegas, NV (pp. 6901–6908).

Yüksel, S., & Başar, T. (2013). *Stochastic networked control systems: Stabilization and optimization under information constraints*. Boston, MA: Birkhäuser.

Zaidi, A. A., Oechtering, T. J., Yüksel, S., & Skoglund, M. (2013). Stabilization and control over Gaussian networks. In G. Como, B. Bernhardsson, & a. Rantzer (Eds.), *Information and control in networks*. Springer.

Serkan Sarıtaş received his B.Sc. degree in Electrical and Electronics Engineering, M.S. degree in Computer Engineering and Ph.D. degree in Electrical and Electronics Engineering from Bilkent University, Turkey in 2010, 2013, and 2018, respectively. Currently he is a Postdoctoral Researcher at the Division of Decision and Control Systems, KTH Royal Institute of Technology. His main research interests include networked control systems, game theory, communication systems, and information theory.

Serdar Yüksel received his B.Sc. degree in Electrical and Electronics Engineering from Bilkent University in 2001; 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 Queen's University as an Assistant Professor of Mathematics and Engineering in the Department of Mathematics and Statistics, where he is now an Associate Professor. He has held a visiting position at Bilkent University. He has been awarded the 2013 CAIMS/PIMS Early Career Award in Applied Mathematics and is currently an associate editor for Automatica and the IEEE Transactions on Automatic Control. His research interests are on stochastic control, information theory, probability, and networked control.

Sinan Gezici received the B.S. degree from Bilkent University, Turkey in 2001, and the Ph.D. degree in Electrical Engineering from Princeton University in 2006. From 2006 to 2007, he worked at Mitsubishi Electric Research Laboratories, Cambridge, MA. 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 was an associate editor for IEEE Transactions on Communications, IEEE Wireless Communications Letters, and Journal of Communications and Networks.