



**HAL**  
open science

## A belief-based approach to signaling

Frédéric Koessler, Marie Laclau, Tristan Tomala

► **To cite this version:**

Frédéric Koessler, Marie Laclau, Tristan Tomala. A belief-based approach to signaling. 2024. halshs-04455227

**HAL Id: halshs-04455227**

**<https://shs.hal.science/halshs-04455227>**

Preprint submitted on 13 Feb 2024

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# A belief-based approach to signaling\*

Frédéric Koessler<sup>†</sup>

Marie Laclau<sup>‡</sup>

Tristan Tomala<sup>§</sup>

February 13, 2024

## Abstract

In this paper, we provide a geometric characterization of the set of interim equilibrium payoff vectors in a general class of signaling games. To obtain a tractable characterization, we use the belief based approach found in the literature on repeated games with incomplete information, cheap talk and Bayesian persuasion. This approach avoids to specify the prior, the strategies of the sender and receiver, and the belief system. The key ingredient is to consider Bayes-plausible belief distributions that are incentive-compatible for the sender. Geometrically, this leads to a constrained convexification of the graphs of the non-revealing interim payoff correspondences. Our characterization extends the analogous result for sender-receiver cheap talk games. We illustrate the results with some classical signaling games. We derive the best equilibrium payoff for the sender when his preferences are type-independent. For zero-sum preferences, we obtain an explicit formula for the ex-ante equilibrium payoff and establish a simple condition for the uniqueness of interim equilibrium payoffs.

KEYWORDS: belief-based approach, cheap talk, information transmission, incomplete information, signaling.

## 1 Introduction

Signaling games represent an archetypical class of strategic and dynamic interactions involving one party possessing private information (the sender) and one uninformed decision-maker (the receiver) observing a costly signal from the sender. These games exhibit their relevance across numerous domains in economics, including instances like job market signaling, where candidates communicate their qualifications to potential employers; reputation analysis, where individuals build and manage their reputations in various social and professional contexts; and within financial markets, where firms signal their financial well-being to entice investors. Additionally, signaling games play a critical role in comprehending advertising strategies within the business realm, where companies use signals to convey product quality to consumers. Beyond economics, signaling games have also found an application in biology, notably exemplified by the handicap principle, where individuals signal their fitness through costly traits.

In this paper, we provide a geometric characterization of (perfect Bayesian) interim equilibrium payoffs for a general class of signaling games. We consider a finite set of types for the sender, but apart from that, we do not make specific assumptions on utility functions, such as single-peakness or single-crossing, and we do not assume that actions, signals, or types are one-dimensional. Within

---

\*The authors gratefully acknowledge the support of the ANR through the program Investissements d'Avenir (ANR-11-IDEX-0003/Labex Ecodec/ANR-11-LABX-0047).

<sup>†</sup>HEC Paris and GREGHEC-CNRS, 1 rue de la Libération, 78351 Jouy-en-Josas, France; koessler@hec.fr

<sup>‡</sup>HEC Paris and GREGHEC-CNRS, 1 rue de la Libération, 78351 Jouy-en-Josas, France; laclau@hec.fr

<sup>§</sup>HEC Paris and GREGHEC, 1 rue de la Libération, 78351 Jouy-en-Josas, France; tomala@hec.fr

this framework, the sender can utilize cheap talk messages in addition to costly signals. We employ a tractable belief-based approach, similar to that used in the literature on repeated games with incomplete information, cheap talk games, and Bayesian persuasion. By using this approach, we eliminate the need to specify the strategies of the sender and receiver, as well as the associated belief system.

Unlike models of cheap talk and Bayesian persuasion, information is conveyed through signals with payoff consequences, so the sender’s interim payoff depends not only on the induced belief of the receiver, but also on the specific signal. Considering the collection of correspondences of non-revealing interim payoffs indexed by the sender’s signal, a *splitting* of the prior belief (i.e., a Bayes-plausible distribution of beliefs) must associate each posterior belief with one of these payoff correspondences. Geometrically, this implies that the set of attainable interim payoff vectors of the sender lies in the convex hull of the graphs of those correspondences. Taking into account the sender’s equilibrium conditions, our method is based on incentive-compatible splittings of the prior type distribution, resulting in a *constrained convexification* of the graphs of the non-revealing payoff correspondences. Notably, our characterization extends the existing equilibrium description for sender-receiver cheap talk games, regardless of any assumptions about the sender’s preferences such as state-independent payoffs.

Our characterization implies that the set of equilibrium payoff vectors does not depend on the set of cheap talk messages, provided it is at least as large as the set of types of the sender: we show that regardless of the sizes of the action, signal and message sets, any equilibrium payoff vector of the sender can be obtained with a sender’s strategy which utilizes no more pairs of signals and messages, than the number of types. We deduce the existence of a perfect Bayesian equilibrium in signaling games with such sets of cheap talk messages. Then, we consider two applications, a simplified poker game and the classical job market signaling game of Spence (1973). Our geometric characterization enables the explicit calculation of all equilibrium payoffs for all possible priors, including for partially revealing equilibria. Finally, we consider two special cases. First, in signaling games with *transparent motives* Lipnowski and Ravid (2020) where the sender’s utility is type independent, we characterize the maximal equilibrium payoff of the sender via the quasi-concave envelope of the best non-revealing sender’s payoff. Second, for zero-sum signaling games, we characterize the ex-ante equilibrium payoff through the concave closure of the maxmin value of the non-revealing game; we show that the interim equilibrium payoff vector is unique when this concave closure is differentiable. Furthermore, in this case, cheap talk messages can be fully dispensed with.

**Related literature** The literature on signaling games, initiated by Spence (1973) in economics and by Zahavi (1975) and Grafen (1990) in biology, has led to significant research in applied work across various fields and in the analysis of equilibrium refinements in games. For a literature review, see, e.g., Kreps and Sobel (1994) and Sobel (2020).

Our belief-based approach to signaling games shares significant connections with the approach employed in the literature on repeated games with incomplete information, cheap talk games, and Bayesian persuasion. Roughly, the idea revolves around viewing the informed player’s strategy as a means to split the prior belief of the uninformed player into a convex combination of posteriors beliefs. These splittings enable the informed player to generate new payoffs within the convex hull of the graphs of the non-revealing payoff correspondences. For a comprehensive review and comparison of how this technique is applied across these areas, refer to Forges (2020).

In the context of zero-sum repeated games where only one player possesses private information, Aumann and Maschler (1966, 1995) and Ponsard and Zamir (1973) showed that the equilibrium payoff for the informed player is given by the concave closure (or concavification) of the minmax value of the one-shot game with no information and common belief, evaluated at the prior. The minmax theorem for those zero-sum games implies that the informed player’s commitment to his strategy is irrelevant.

The approach can be applied to one-shot non-zero-sum scenarios when the informed player is able

to commit to an information disclosure strategy. For instance, in the Bayesian persuasion setting introduced by Kamenica and Gentzkow (2011), the sender commits to his messaging strategy before learning the state. The maximum ex-ante expected payoff he can obtain is given by the concave closure  $\text{cav } w(p_0)$ , where  $w(p)$  is the highest ex-ante expected payoff for the sender when the uninformed receiver selects an optimal action based on belief  $p$ .

Hart (1985) expands upon Aumann and Maschler’s analysis to encompass non-zero-sum repeated games and provides a geometric characterization of all Nash equilibrium payoffs in undiscounted games. Since no commitment is assumed, the characterization needs to account for the informed player’s incentive-compatible conditions, and it keeps track of the vector of interim payoffs for each type of the informed player. These ideas have been adapted and extended by Forges (1984, 1990) and Aumann and Hart (2003) for characterizing equilibrium payoffs in cheap talk games where both players send messages over multiple periods before taking action with payoff consequences. Similar techniques have been employed by Chakraborty and Harbaugh (2010) and Lipnowski and Ravid (2020) to characterize equilibrium payoffs in cheap talk games with transparent motives.

Unlike the cheap talk literature, our focus centers on signals that are payoff-relevant either for the sender, for the receiver, or both. Technically, this implies that the sender can split the prior belief and the corresponding payoff vector on multiple graphs of non-revealing payoff correspondences, one for each signal. In cases where all signals of the sender are payoff-irrelevant, the characterization reduces to the one mentioned above for sender-receiver cheap talk games. Our characterization is also related to the one for disclosure games with hard information in Forges and Koessler (2005), by considering signals that are payoff-irrelevant for the receiver, and either payoff-irrelevant or very costly for the sender, depending on his type.

Lastly, our characterization is related to the recent contribution by Boleslavsky and Shadmehr (2023), who extend the Bayesian persuasion model to include payoff-relevant signals. They characterize the maximal ex-ante expected payoff the sender can achieve in signaling games in which the sender ex-ante commit to his signaling strategy. Like our work, their approach revolves around the graphs of non-revealing payoff correspondences for each signal. However, because we do not assume commitment, in line with the cheap talk literature we keep track of interim payoff vectors and maintain incentive-compatible conditions for the sender.

## 2 Model

**Setup** There are two players, a sender and a receiver. The nonempty and finite set of types of the sender is denoted by  $T$ . The sender is privately informed of his type, while the receiver is uninformed. The prior probability distribution is denoted by  $p_0 \in \Delta(T)$  and is assumed to have full support. The set of actions of the receiver is denoted by  $A$ .<sup>1</sup> The sender has a set of costly signals  $S$  and a set of costless (cheap talk) messages  $M$ . We assume that  $A$ ,  $S$ , and  $M$  are nonempty compact metric spaces. Except stated otherwise, we assume that  $M$  has at least  $|T|$  elements. Given an action  $a \in A$ , a signal  $s \in S$ , and a type  $t \in T$ , the utility of the sender is  $u(a, s, t)$ , and the utility of the receiver is  $u_R(a, s, t)$ . The utilities  $u(a, s, t)$  and  $u_R(a, s, t)$  are continuous on  $A \times S \times T$ . Utility functions are extended to the set of mixed actions  $\Delta(A)$ <sup>2</sup> in the usual way.

**Signaling game** The timing of the signaling game is as follows:

1. The sender’s type  $t$  is drawn from  $T$  according to  $p_0$ .

---

<sup>1</sup>We can allow the set of actions of the receiver to depend on the signal, i.e., let  $A(s)$  denote the set of actions of the receiver given  $s$ .

<sup>2</sup>For a metric compact  $X$ ,  $\Delta(X)$  denotes the set of Borel probability measures over  $X$ .

2. The sender observes  $t \in T$ , chooses a signal  $s \in S$  and a cheap talk message  $m \in M$ .
3. The receiver observes the signal  $s$  and the cheap talk message  $m$ , then chooses an action  $a \in A$ .

A strategy for the sender is  $\sigma : T \rightarrow \Delta(S \times M)$ , and a strategy for the receiver is a Borel measurable  $\tau : S \times M \rightarrow \Delta(A)$ . The interim payoff vector for the sender induced by the strategy profile  $(\sigma, \tau)$  is given by  $v = (v_t)_{t \in T}$  with

$$v_t = \int_{S \times M} u(\tau(s, m), s, t) d\sigma(s, m | t), \text{ for every } t \in T.^3$$

A belief system is given by a Borel measurable  $\mu : S \times M \rightarrow \Delta(T)$ , where  $\mu(t | s, m)$  is the belief assigned by the receiver to type  $t$  when he observes  $(s, m)$ .

**Perfect Bayesian equilibrium** For every  $p \in \Delta(T)$  and  $s \in S$ , let  $Y(s, p)$  be the set of mixed actions that are optimal for the receiver given signal  $s$  and belief  $p$ . That is:

$$Y(s, p) := \arg \max_{y \in \Delta(A)} u_R(y, s, p),$$

where, with some abuse of notation,  $u_R(y, s, p) = \sum_t p(t) u_R(y, s, t) = \sum_t p(t) \int_A u_R(a, s, t) dy(a)$ .

A strategy profile  $(\sigma, \tau)$ , with payoff vector  $v = (v_t)_{t \in T}$ , is a perfect Bayesian equilibrium (PBE) of the signaling game if there exists a belief system  $\mu$  such that the following three conditions are satisfied. First, for each type, the sender cannot increase his interim expected payoff by deviating to any pair of signal and message. Second, for every pair of signal and message, the receiver chooses optimal actions given the signal and his belief about the sender's type. Finally, the receiver's belief is computed using Bayes' rule whenever possible. Formally:

- (i) Sequential rationality for the sender: for every  $t \in T$  and  $(s, m) \in S \times M$ ,

$$v_t \geq u(\tau(s, m), s, t);$$

- (ii) Sequential rationality for the receiver: for every  $(s, m) \in S \times M$ ,

$$\tau(s, m) \in Y(s, \mu(s, m));$$

- (iii) Belief consistency: for every  $t \in T$  and every Borel set  $B \subseteq S \times M$ ,

$$p_0(t) \sigma(B | t) = \sum_{\tilde{t} \in T} p_0(\tilde{t}) \int_B \mu(t | s, m) d\sigma(s, m | \tilde{t}).$$

Such a triple  $(\sigma, \tau, \mu)$  is called a PBE assessment. We say that a payoff vector for the sender  $v \in \mathbb{R}^T$  is a *PBE payoff vector* if it is induced by some PBE strategy profile.

---

<sup>3</sup>We focus on the interim payoff vectors of the sender for simplicity, the expected payoff of the receiver can be easily included in the analysis.

### 3 Geometric characterization of equilibrium payoffs

For every  $s \in S$  and  $p \in \Delta(T)$ , let  $\mathcal{E}(s, p)$  be the set of all feasible vectors of interim payoffs of the sender when the receiver chooses an optimal (mixed) action given  $s$  and  $p$ :

$$\mathcal{E}(s, p) := \{v \in \mathbb{R}^T : \exists y \in Y(s, p) \text{ s.t. } v = (u(y, s, t))_t\}.$$

We call this the set of *non-revealing payoffs* for the sender at  $p$  given  $s$ .

Observe that  $\mathcal{E}(s, p)$  depends on the signal  $s$  through two distinct channels: through  $Y(s, p)$ , when the receiver's optimal action varies with  $s$  for a given posterior  $p$ , and through the direct impact of  $s$  on the utility of the sender for a given action of the receiver. We say that the signal  $s$  is cheap talk if it does not affect the utilities of the sender and of the receiver; in this case,  $\mathcal{E}(s, p)$  is independent of  $s$ , as in Forges (1984, 1990, 1994) and Aumann and Hart (2003).

We define then the *modified* set of non-revealing payoffs for the sender at  $p$  given  $s$  as:

$$\mathcal{E}^+(s, p) := \{v \in \mathbb{R}^T : \exists y \in Y(s, p) \text{ s.t. } \forall t, v_t \geq u(y, s, t) \text{ and } v_t = u(y, s, t) \text{ if } p(t) > 0\}.$$

To obtain the modified non-revealing payoff correspondence, consider for any  $p$  in  $\Delta(T)$ , the non-revealing payoffs given the signal  $s$  and, at the boundary of the simplex, adjust these payoffs by granting higher payoffs to the sender's types with vanishing probabilities. Notice that  $\mathcal{E}(s, p) \subseteq \mathcal{E}^+(s, p)$ , and  $\mathcal{E}(s, p) = \mathcal{E}^+(s, p)$  if  $p$  has full support.

The graph of the modified non-revealing payoff correspondence of the sender given  $s$  is denoted by:

$$G_s := \text{gr } \mathcal{E}^+(s, \cdot) := \{(v, p) \in \mathbb{R}^T \times \Delta(T) : v \in \mathcal{E}^+(s, p)\}.$$

We denote by  $G = \bigcup_{s \in S} G_s$  the union of the graphs of the non-revealing payoff correspondences for all  $s$ , that is,

$$G = \{(v, p) \in \mathbb{R}^T \times \Delta(T) : \exists s \in S, y \in Y(s, p) \text{ s.t. } \forall t, v_t \geq u(y, s, t) \text{ and } v_t = u(y, s, t) \text{ if } p(t) > 0\}.$$

Let  $\text{co}_v(G)$  be the set of payoff vectors obtained by convexifying  $G$  with respect to  $p$ , when the vector payoff  $v$  is kept constant. Formally this is:

$$\text{co}_v(G) := \left\{ \left( v, \sum_{k=1}^{|T|} \lambda_k p_k \right) \in \mathbb{R}^T \times \Delta(T) : (\lambda_k)_{k=1}^{|T|} \in \Delta(\{1, \dots, |T|\}), (v, p_k) \in G \forall k \right\}.$$

Since  $p$  lies in the  $(|T| - 1)$ -dimensional simplex, only convex combination of cardinality at most  $|T|$  are considered, thanks to Carathéodory's theorem. A family  $(\lambda_k, p_k)_{k=1}^{|T|}$  such that  $p_0 = \sum_{k=1}^{|T|} \lambda_k p_k$  and  $(v, p_k)$  belongs to  $G$  for all  $k$ , will be called an *incentive compatible splitting* of  $p_0$ .

Next, we introduce an individual rationality condition. Let  $\text{INTIR}_s$  be the set of all payoff vectors  $v \in \mathbb{R}^T$  such that there exists  $p \in \Delta(T)$  and  $y \in Y(s, p)$  satisfying  $v_t \geq u(y, s, t)$  for every  $t \in T$ . Finally, let

$$\text{INTIR} := \bigcap_{s \in S} \text{INTIR}_s,$$

be the set of all payoff vectors  $v \in \mathbb{R}^T$  such that for every  $s \in S$  there exists  $p \in \Delta(T)$  and  $y \in Y(s, p)$  such that  $v_t \geq u(y, s, t)$  for every  $t \in T$ . The payoff vectors in  $\text{INTIR}$  are called interim individually rational.<sup>4</sup> Observe that  $\text{INTIR}$  is directly obtained from  $G$ , this is the set of payoff vectors which are coordinate-wise above some point in the projection of  $G$  on the set of payoff vectors  $\mathbb{R}^T$ :

$$\text{INTIR} = \text{proj}_{\mathbb{R}^T}(G) + \mathbb{R}_+^T.$$

---

<sup>4</sup>This individual rationality condition requires the receiver to choose a mixed action which is optimal for the receiver for some belief. A weaker condition requires that for every  $s \in S$ , there exists  $y \in \Delta(A)$  such that  $v_t \geq u(y, s, t)$  for every  $t \in T$ . This weaker form is the relevant condition for the study of Nash equilibria instead of PBE of the signaling game.

Our main result is that the interim individual rationality condition INTIR and the constrained convex hull of modified non-revealing payoffs  $\text{co}_v(G)$ , fully pin down the set of PBE payoff vectors. This is stated in the following theorem.

**Theorem 1.** *The payoff vector  $v \in \mathbb{R}^T$  is a PBE payoff vector of the sender in the signaling game if and only if  $v \in \text{INTIR}$  and  $(v, p_0) \in \text{co}_v(G)$ .*

*Proof.*

(i) *From constrained convexification to equilibrium.*

Let  $v \in \text{INTIR}$  be such that  $(v, p_0) \in \text{co}_v(G)$ , we construct a PBE assessment  $(\sigma, \tau, \mu)$  with payoff vector  $v$ , where  $\sigma$  has finite support. Since  $(v, p_0) \in \text{co}_v(G)$ , there exists  $K \in \{1, \dots, |T|\}$ ,  $(\lambda_k)_{k=1}^K \in \Delta(\{1, \dots, K\})$ ,  $\lambda_k > 0$ ,  $s_k \in S$  and  $y_k \in Y(s_k, p_k)$  for every  $k$  such that:

$$p_0 = \sum_{k=1}^K \lambda_k p_k, \quad (1)$$

and for every  $k = 1, \dots, K$ :

$$v_t \geq u(y_k, s_k, t) \text{ and } v_t = u(y_k, s_k, t) \text{ if } p_k(t) > 0. \quad (2)$$

Consider a profile of messages  $(m_k)_k$  with  $m_k \in M$  and  $m_k \neq m_{k'}$  for every  $k \neq k'$ , this is possible as  $|M| \geq |T| \geq K$ . Since  $v \in \text{INTIR}$ , for every  $s \in S$ , there exists  $q_s \in \Delta(T)$  and

$$\bar{y}_s \in Y(q_s, s), \quad (3)$$

such that:

$$v_t \geq u(\bar{y}_s, s, t), \text{ for every } t \in T. \quad (4)$$

Consider the following strategies  $\sigma$  for the sender and  $\tau$  for the receiver:

$$\sigma(s_k, m_k | t) = \frac{p_k(t)\lambda_k}{p_0(t)}, \text{ for every } k, \quad (5)$$

$$\sigma(s, m | t) = 0, \text{ if } (s, m) \neq (s_k, m_k) \text{ for every } k, \quad (6)$$

$$\tau(s, m) = \begin{cases} y_k & \text{if } (s, m) = (s_k, m_k) \text{ for some } k, \\ \bar{y}^s & \text{otherwise.} \end{cases} \quad (7)$$

Consider the following belief system:

$$\mu(t | s, m) = \begin{cases} p_k & \text{if } (s, m) = (s_k, m_k) \text{ for some } k, \\ q_s & \text{otherwise.} \end{cases} \quad (8)$$

For the sender of type  $t$ , the interim payoff induced by  $(\sigma, \tau)$  is indeed  $v_t$  because:

$$\begin{aligned} \int_{S \times M} u(\tau(s, m), s, t) d\sigma(s, m | t) &= \sum_{k=1}^K \frac{p_k(t)\lambda_k}{p_0(t)} u(y_k, s_k, t), \text{ by (5) and (7),} \\ &= v_t, \text{ by (2) and (1).} \end{aligned}$$

Then, sequential rationality for the sender (PBE condition (i)) follows directly from (2), (4), and (7), sequential rationality for the receiver (PBE condition (ii)) follows directly from (3), (7),

and (8). Finally, to verify belief consistency (PBE condition (iii)), it remains to show that the probability of  $t$  conditional on  $(s_k, m_k)$ , given by (8), is equal to  $p_k(t)$  for every  $k$  and  $t$ . From Bayes' rule and (5), this conditional probability is:

$$\frac{\sigma(s_k, m_k | t) p_0(t)}{\sum_{\tilde{t} \in T} \sigma(s_k, m_k | \tilde{t}) p_0(\tilde{t})} = \frac{\frac{p_k(t) \lambda_k}{p_0(t)} p_0(t)}{\sum_{\tilde{t} \in T} \frac{p_k(\tilde{t}) \lambda_k}{p_0(\tilde{t})} p_0(\tilde{t})} = \frac{p_k(t) \lambda_k}{\sum_{\tilde{t} \in T} p_k(\tilde{t}) \lambda_k} = p_k(t).$$

This completes the “if” part of the theorem.

(ii) *From equilibrium to constrained convexification.*

Consider a PBE assessment  $(\sigma, \tau, \mu)$  with payoff vector  $v$ . First, we show that  $v \in \text{INTIR}$ . For every  $s \in S$ , consider an arbitrary pair of signal and message  $(s, m) \in S \times M$  and let  $y = \tau(s, m)$ . Since  $\tau$  is sequentially rational for the receiver, we have  $y \in Y(s, p)$  for  $p = \mu(s, m) \in \Delta(T)$ . Since  $\sigma$  is sequentially rational for the sender, we have  $v_t \geq u(y, s, t)$  for every  $t$ . We conclude that for every signal  $s$ , we have  $v \in \text{INTIR}_s$ , therefore  $v \in \text{INTIR}$ .

Second, we show that  $(v, p_0) \in \text{co}_v(G)$ . Let  $\lambda(\cdot) := \sum_{t \in T} p_0(t) \sigma(\cdot | t) \in \Delta(S \times M)$  denote the marginal distribution on  $S \times M$  induced by  $p_0$  and  $\sigma$ . To prove that  $(v, p_0) \in \text{co}_v(G)$ , it suffices to show that there exists  $\tilde{X} \subseteq S \times M$  with  $\lambda(\tilde{X}) = 1$  such that (a)  $v \in \mathcal{E}^+(s, \mu(s, m))$  for every  $(s, m) \in \tilde{X}$ , and (b)  $p_0 \in \text{co}(\{\mu(s, m) : (s, m) \in \tilde{X}\})$ .

The PBE condition for the sender:

$$v_t \geq u(\tau(s, m), s, t) \quad \forall (s, m) \in S \times M, \quad (9)$$

implies:

$$\sum_{t \in T} (v_t - u(\tau(s, m), s, t)) \mu(t | s, m) \geq 0, \quad \text{for every } (s, m) \in S \times M. \quad (10)$$

We also have:

$$\begin{aligned} & \mathbb{E}_\lambda \left[ \sum_t (v_t - u(\tau(s, m), s, t)) \mu(t | s, m) \right] \\ &= \int_{S \times M} \sum_t (v_t - u(\tau(s, m), s, t)) \mu(t | s, m) d\lambda(s, m) \\ &= \sum_t \int_{S \times M} (v_t - u(\tau(s, m), s, t)) \mu(t | s, m) d\lambda(s, m) \\ &= \sum_t \int_{S \times M} (v_t - u(\tau(s, m), s, t)) p_0(t) d\sigma(s, m | t) \\ &= \sum_t v_t p_0(t) \underbrace{\int_{S \times M} d\sigma(s, m | t)}_{=1} - \sum_t p_0(t) \underbrace{\int_{S \times M} u(\tau(s, m), s, t) d\sigma(s, m | t)}_{=v_t} = 0. \end{aligned}$$

Hence, combined with (10), the previous equality implies that,  $\lambda$ -almost surely:

$$\sum_t (v_t - u(\tau(s, m), s, t)) \mu(t | s, m) = 0.$$

Thus, from (9), there exists a Borel subset  $\tilde{X} \subseteq S \times M$  with  $\lambda(\tilde{X}) = 1$  such that:

$$v_t = u(\tau(s, m), s, t), \quad \text{for all } (s, m) \in \tilde{X} \text{ and } t \in T \text{ s.t. } \mu(t | s, m) > 0. \quad (11)$$



The PBE condition for the receiver implies:

$$\tau(s, m) \in Y(s, \mu(s, m)), \text{ for all } (s, m) \in \tilde{X}. \quad (12)$$

Putting (9), (11), and (12) together, we get  $v \in \mathcal{E}^+(s, \mu(s, m))$  for every  $(s, m) \in \tilde{X}$ , which proves (a).

Now, from belief consistency with  $B = \tilde{X}$  we have:

$$p_0 = \int_{\tilde{X}} \mu(s, m) d\lambda(s, m).$$

Hence, property (b) directly follows from the following lemma.

**Lemma 1.** *Let  $p_0 \in \Delta(T)$ ,  $\sigma : T \rightarrow \Delta(S \times M)$ , and denote by  $\lambda(\cdot) := \sum_{t \in T} p_0(t) \sigma(\cdot | t) \in \Delta(S \times M)$  the marginal distribution on  $S \times M$  induced by  $p_0$  and  $\sigma$ . Let  $\tilde{X}$  be a Borel subset of  $S \times M$  such that  $\lambda(\tilde{X}) = 1$ , and let  $\mu : S \times M \rightarrow \Delta(T)$  be a Borel measurable function satisfying:*

$$p_0(t) = \int_{\tilde{X}} \mu(t | x) d\lambda(x). \quad (13)$$

Then,

$$p_0 \in \text{co}(\{\mu(x) : x \in \tilde{X}\}).$$

Notice that the conclusion that  $p_0$  belongs to the convex hull of the beliefs  $\{\mu(x) : x \in \tilde{X}\}$ , is stronger than requiring that it merely belongs to the closure of the convex hull. *Proof.* Let  $Z \subseteq \tilde{X}$  be a Borel subset of  $\tilde{X}$  such that  $\lambda(Z) = 1$  and such that the dimension  $d(Z)$  of the convex hull of  $\{\mu(x) : x \in Z\}$  is minimal among the  $d(Z')$ , for all Borel subsets  $Z' \subseteq \tilde{X}$  with  $\lambda(Z') = 1$ . We have:

$$\forall t \in T, p_0(t) = \int_Z \mu(t|x) d\lambda(x).$$

Suppose that  $p_0 \notin \text{co}(\{\mu(x) : x \in Z\})$ . From the separation theorem, there exists  $\psi \in \mathbb{R}^T$  such that:

$$\forall x \in Z, \sum_t p_0(t) \psi_t \geq \sum_t \mu(t|x) \psi_t,$$

where the separating vector  $\psi$  is not 0 and belongs to the  $d(Z)$ -dimensional affine hull of  $\{\mu(x) : x \in Z\}$ . Integrating this inequality gives:

$$\sum_t p_0(t) \psi_t \geq \int_Z \sum_t \mu(t|x) \psi_t d\lambda(x) = \sum_t \int_Z \mu(t|x) d\lambda(x) \psi_t = \sum_t p_0(t) \psi_t.$$

It follows that  $\sum_t \mu(t|x) \psi_t = \sum_t p_0(t) \psi_t$ ,  $\lambda$ -almost surely, that is:

$$\lambda\left(\left\{x \in Z : \sum_t \mu(t|x) \psi_t = \sum_t p_0(t) \psi_t\right\}\right) = 1.$$

Therefore,  $Z' = \{x \in Z : \sum_t \mu(t|x) \psi_t = \sum_t p_0(t) \psi_t\}$  satisfies  $\lambda(Z') = 1$  and  $d(Z') < d(Z)$  which is a contradiction. We conclude that  $p_0 \in \text{co}(\{\mu(x) : x \in Z\})$  and therefore  $p_0 \in \text{co}(\{\mu(x) : x \in \tilde{X}\})$ . ■

This completes the “only if” part of the theorem and ends the proof. ■

**Consequences** As illustrated in Section 4, this geometric characterization enables the explicit calculation of all equilibrium payoffs, for all possible priors, in various examples and applications. Furthermore, it has several interesting consequences. Corollary 1 illustrates a first direct consequence of Theorem 1: to characterize all PBE payoff vectors of the signaling game with enough cheap talk messages, it suffices to focus on the sender's strategies with finite support, of cardinality at most  $|T|$ , regardless of the sizes of the action, signal and message spaces.<sup>5</sup> This result also implies that the set of PBE payoff vectors does not depend on the message space  $M$ , provided that  $M$  contains at least  $|T|$  elements.

**Corollary 1.** *The payoff vector  $v \in \mathbb{R}^T$  is a PBE payoff vector of the sender in the signaling game if and only if  $v$  is a payoff vector of a PBE in which at most  $|T|$  pairs of signal and message are used with positive probability.*

What about the existence of PBE? If the game is finite (i.e., both sets of signals  $S$  and of actions  $A$  are finite), then there exists at least one sequential equilibrium (Kreps and Wilson, 1982, Proposition 1), and any sequential equilibrium is a PBE. Under the compactness and continuity assumptions of our model, Manelli (1996) shows that the signaling game has a PBE as long as the message space is large enough, precisely for  $M = \Delta(A)$ .<sup>6</sup> Corollary 1 shows that for any such PBE, the same payoff vector can be obtained through a PBE of the signaling game with any message set of cardinality at least  $|T|$ . Hence, due to our maintained assumption that  $M$  contains at least  $|T|$  elements, Corollary 2 in Manelli (1996) and Corollary 1 imply the existence of a PBE in our signaling game. Summarizing:

**Corollary 2.** *The signaling game has at least one PBE and the set of payoff vectors:*

$$\{v \in \mathbb{R}^T : v \in \text{INTIR} \text{ and } (v, p_0) \in \underset{v}{\text{co}}(G)\},$$

*is non-empty.*

A cheap talk game is a signaling game where signals are payoff-irrelevant, i.e.,  $u(a, s, t) = u(a, s', t)$  and  $u_R(a, s, t) = u_R(a, s', t)$  for all signals  $s$  and  $s'$ . Theorem 1 characterizes PBE payoff vectors for a cheap talk game as a particular case. In such a game, we have  $G = G_s$  for every  $s$ , and the condition  $v \in \text{INTIR}$  is irrelevant as it is straightforwardly implied by  $(v, p_0) \in \text{co}_v(G)$ . Therefore, we have the following corollary of Theorem 1.

**Corollary 3.** *Assume that signals are payoff-irrelevant. The payoff vector  $v \in \mathbb{R}^T$  is a PBE payoff vector of the sender in the cheap talk game iff  $(v, p_0) \in \text{co}_v(G)$ .*

Corollary 3 is well known in the cheap talk literature.<sup>7</sup> Notice that in a cheap talk game, every non-revealing payoff vector  $v \in \mathcal{E}(s, p_0)$  is a “babbling” PBE payoff vector. However, in a signaling game with payoff-relevant messages, non-revealing payoff vectors may fail to be interim individually rational and thus a non-revealing PBE may not exist. That is, in a signaling game, the set  $\mathcal{E}(s, p_0) \cap \text{INTIR}$  may be empty for every  $s$  (see Section 4 for some examples). We deduce the following.

**Corollary 4.** *The payoff vector  $v$  is a non-revealing PBE payoff vector if and only if there exists a signal  $s$  such that  $v \in \mathcal{E}(s, p_0) \cap \text{INTIR}$ . Thus, the signaling game has a non-revealing PBE if and only if  $(\cup_s \mathcal{E}(s, p_0)) \cap \text{INTIR} \neq \emptyset$ .*

<sup>5</sup>Taking into account the receiver's ex ante expected payoff would require  $|T| + 1$  pairs of signal and message instead of  $|T|$ .

<sup>6</sup>A PBE may not exist for continuous sets of signals if the message space is too small, see the examples in Manelli (1996).

<sup>7</sup>The characterization of Corollary 3 is known for Nash equilibria of cheap talk games where PBE and Nash equilibrium payoffs coincide.

## 4 Examples

### 4.1 Simple poker

Two risk-neutral individuals are engaged in a simplified poker game as follows. Initially, both players place 1 euro each on the table as the blind. Player 1 (the sender) randomly draws a card, which can be either high ( $t_H$ ) or low ( $t_L$ ), with probabilities  $p_0(t_H) = p_0$  and  $p_0(t_L) = 1 - p_0$  respectively. Only the sender has knowledge of the drawn card. Then, the sender chooses between multiple available signals: he can either fold (signal  $s = 0$ ), in which case player 2 (the receiver) earns the 2 euros on the table, thereby concluding the game. Alternatively, the sender can opt to raise  $s$  times the blind, with  $s \in \{1, 2, \dots, \bar{s}\}$ , necessitating an additional bet of  $s$  euros. The raise is bounded by  $s = \bar{s}$ , corresponding to the “all in” scenario. Finally, it is the receiver’s turn, who must choose between two options: he can fold (action  $a = f$ ), causing the sender to claim the  $2 + s$  euros on the table; or the receiver can call (action  $a = c$ ), which involves placing an additional bet of  $s$  euros to reveal the sender’s card. If the card is high, the sender secures the  $2 + 2s$  euros on the table; whereas if the card is low, the receiver claims the winnings.

Hence, the sender’s (net) payoff as a function of the state  $t \in T = \{t_L, t_H\}$ , the signal  $s \in S = \{0, 1, 2, \dots, \bar{s}\}$ , and the receiver’s action  $a \in \{f, c\}$ , is given by:

$$u(a, 0, t) = -1,$$

$$u(f, s, t) = 1, \quad u(c, s, t_H) = 1 + s, \quad \text{and} \quad u(c, s, t_L) = -1 - s, \quad \text{for } s > 0.$$

The payoff for the receiver is  $u_R(a, s, t) = -u(a, s, t)$ . For  $s > 0$ , the set of optimal mixed actions for the receiver as a function of  $p = p(t_H) \in [0, 1]$  is given by:

$$Y(s, p) = \begin{cases} \{c\} & \text{if } p < \frac{2+s}{2+2s}, \\ \Delta(\{c, f\}) & \text{if } p = \frac{2+s}{2+2s}, \\ \{f\} & \text{if } p > \frac{2+s}{2+2s}. \end{cases}$$

Observe that  $\bar{p}(s) := \frac{2+s}{2+2s}$  is strictly decreasing in  $s$ , with  $\bar{p}(1) = \frac{3}{4}$  and  $\lim_{s \rightarrow \infty} \bar{p}(s) = \frac{1}{2}$ . For  $s = 0$ , the action of the receiver is irrelevant:  $Y(0, p) = \Delta(\{c, f\})$  for every  $p$ .

The set of non-revealing payoffs for the sender at  $p$  given  $s > 0$  is given by:

$$\mathcal{E}(s, p) = \begin{cases} \{(1 + s, -1 - s)\} & \text{if } p < \bar{p}(s), \\ \text{co}\{(1 + s, -1 - s), (1, 1)\} & \text{if } p = \bar{p}(s), \\ \{(1, 1)\} & \text{if } p > \bar{p}(s), \end{cases}$$

and for  $s = 0$ ,

$$\mathcal{E}(0, p) = \{(-1, -1)\}.$$

The interim individually rational payoff vectors are given by

$$\text{INTIR}_0 = \{(v_H, v_L) : v_H \geq -1, v_L \geq -1\},$$

$$\text{INTIR}_s = \left\{ (v_H, v_L) : v_H \geq 1, v_L \geq -1 - s, v_L \geq \frac{-2 - s}{s}v_H + \frac{2 + 2s}{s} \right\}.$$

Hence:

$$\text{INTIR} = \bigcap_{s=0}^{\bar{s}} \text{INTIR}_s = \left\{ (v_H, v_L) : v_H \geq 1, v_L \geq -1, v_L \geq \frac{-2 - \bar{s}}{\bar{s}}v_H + \frac{2 + 2\bar{s}}{\bar{s}} \right\}.$$

Figure 1 represents the projections of the graphs of  $\mathcal{E}^+(s, \cdot)$  on the sender's vector payoffs for three possible signals:  $s = 0$  (in green),  $s = 2$  (in red), and  $s = \bar{s} = 4$  (in blue). The points at the north-east of these curves satisfy the INTIR condition, represented by the gray area. In the INTIR set, the curves intersect at  $(v_H, v_L) = (1, 1)$  and  $(v_H, v_L) = (\frac{7}{3}, -1)$ . The unique non-revealing payoff that belongs to INTIR is then  $(v_H, v_L) = (1, 1)$ , for every  $s > 0$  and  $p_0 > \bar{p}(s) = \frac{2+s}{2+2s}$ . Indeed, the non-revealing payoff  $(v_H, v_L) = (-1, -1)$  for  $s = 0$  and  $p_0 \in (0, 1)$  does not belong to INTIR, so it is never a PBE payoff. Similarly, the non-revealing payoffs  $(v_H, v_L) = (1 + s, -1 - s)$  for  $s > 0$  and  $p_0 < \bar{p}(s) = \frac{2+s}{2+2s}$  do not belong to INTIR, so they are never PBE payoffs.

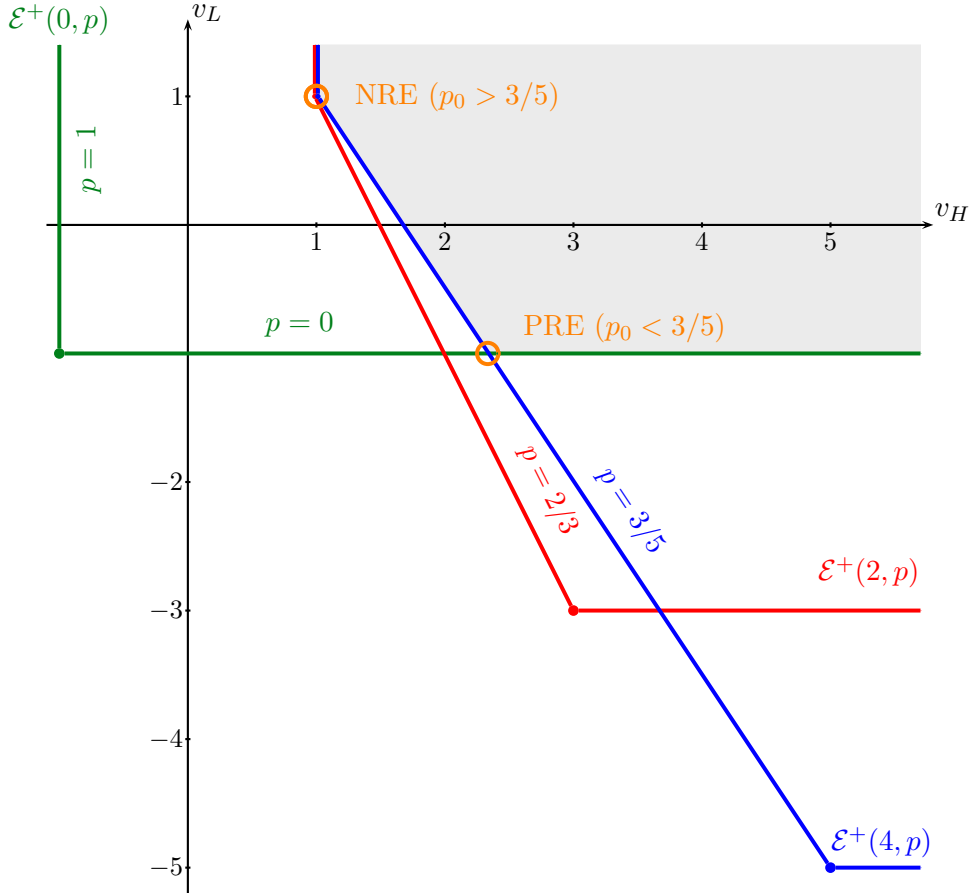


Figure 1: Graphs of  $\mathcal{E}^+(s, \cdot)$  (projected on the space of the sender vector payoffs) in the simplified poker for  $s = 0$  (in green),  $s = 2$  (in red), and  $s = 4$  (in blue).

Then,  $(v_H, v_L) = (\frac{7}{3}, -1)$  is the only candidate for an incentive compatible splitting. Such a partially revealing splitting, with beliefs  $p' = 0$  (on the green curve) and  $p'' = \bar{p}(4) = \frac{3}{5}$  (on the blue curve), is feasible for all priors  $p_0 \in (0, \frac{3}{5})$ . In this equilibrium, the high type always raises  $s = \bar{s} = 4$  euros, and the low type randomizes between folding ( $s = 0$ ) and raising  $s = 4$  euros in such a way that the receiver's belief after seeing a raise of  $s = 4$  euros is  $p = \frac{3}{5}$  (so the low type chooses  $s = 4$  with probability  $\frac{2p_0}{3(1-p_0)}$ , and  $s = 0$  with the complementary probability).

In this example where the raise is bounded by  $\bar{s} = 4$  euros, we conclude that the unique PBE payoff is  $(v_H, v_L) = (\frac{7}{3}, -1)$  if  $p_0 < \frac{3}{5}$  (represented by the orange circle PRE in the figure), and the unique PBE payoff is  $(v_H, v_L) = (1, 1)$  if  $p_0 > \frac{3}{5}$  (represented by the orange circle NRE in the figure).

The geometric characterization of equilibria easily extends to any bound  $\bar{s}$ . If  $p_0 > \bar{p}(\bar{s}) = \frac{2+\bar{s}}{2+2\bar{s}}$ ,

every PBE is non-revealing: the sender raises with the same amount  $s > 0$ , where  $p_0 > \bar{p}(s) = \frac{2+s}{2+2s}$  regardless of his type. The receiver always folds, and the unique PBE payoff is  $(v_H, v_L) = (1, 1)$ . If  $p_0 < \bar{p}(\bar{s}) = \frac{2+\bar{s}}{2+2\bar{s}}$ , the PBE outcome is partially revealing and unique: the sender splits the prior  $p_0$  into beliefs  $p' = 0$  (with  $s = 0$ ) and  $p'' = \frac{2+\bar{s}}{2+2\bar{s}}$  (with  $s = \bar{s}$ ). The sender raises  $\bar{s}$  euros when  $t = t_H$ , and randomizes between  $s = \bar{s}$  (with probability  $\frac{p_0\bar{s}}{(1-p_0)(2+\bar{s})}$ ) and  $s = 0$  (with the complementary probability) when  $t = t_L$ . The corresponding payoff vector for the sender is  $(\frac{2+3\bar{s}}{2+\bar{s}}, -1)$ . Notice that  $\frac{2+3\bar{s}}{2+\bar{s}}$  is strictly increasing in  $\bar{s}$ , so the value of the game is strictly increasing in  $\bar{s}$  (and it only benefits the sender of type  $t_H$ ), but it is bounded, with  $\lim_{\bar{s} \rightarrow \infty} \frac{2+3\bar{s}}{2+\bar{s}} = 3$ .

Therefore, in this simplified poker when  $p_0 < \bar{p}(\bar{s}) = \frac{2+\bar{s}}{2+2\bar{s}}$  (and in particular, for  $p_0 = \frac{1}{2}$ ), it is never optimal to raise with an intermediate bet  $s \in (0, \bar{s})$ : the informed player either folds ( $s = 0$ ) or goes “all in” ( $s = \bar{s}$ ). The expected payoff of the informed player is strictly increasing in the amount  $\bar{s}$  of chips the player has in his stack. In this example, the PBE outcomes are the same as the PBE outcomes of the signaling game without cheap talk.

## 4.2 Spence job market signaling

Consider the classical textbook job market signaling game (e.g., Mas-Colell, Whinston, and Green, 1995) where the sender is a worker with two possible types,  $T = \{t_L, t_H\}$ . Type  $t \in T$  corresponds to the worker’s productivity and  $t_H > t_L \geq 0$ . With some abuse of notation, we identify a probability distribution over  $T$  by the corresponding probability of  $t_H$ , so  $p_0 = p_0(t_H) \in (0, 1)$  denotes the prior probability that the worker’s productivity is high. A signal  $s \geq 0$  is the amount of education obtained by the worker. The cost of obtaining education level  $s$  for a type  $t$  worker is given by a twice continuously differentiable function  $c(s, t)$ , where  $c(0, t) = 0$ ,  $c_s(s, t) > 0$ ,  $c_t(s, t) < 0$  for  $s > 0$ ,  $c_{ss}(s, t) > 0$  and  $c_{st}(s, t) < 0$ .

The receiver is an employer in a competitive market. His action is a wage offer  $y \geq 0$ , and let’s assume that a sequentially rational wage offer given belief  $p = p(t_H)$  is simply the expected productivity of the worker,  $y(p) = pt_H + (1-p)t_L$ . Using the notations of our model, this can be obtained from the utility function of the receiver  $u_R(y, s, t) = -(y-t)^2$ , whereby  $Y(s, p) = \{y(p)\}$  for every  $p$  and  $s$ . The utility function of the sender is given by:

$$u(y, s, t) = f(y) - c(s, t),$$

where  $f(y)$  is continuous and strictly increasing. Without loss of generality, we normalize  $t_H$  and  $f(1)$  to 1, and  $t_L$  and  $f(0)$  to 0, so  $y(p) = p$  and  $f(y(p)) = f(p)$  is strictly increasing from 0 to 1.

It is immediate to show that  $\text{INTIR} = \mathbb{R}_+^2$ ,  $\mathcal{E}(s, p) = \{(f(p) - c(s, t_H), f(p) - c(s, t_L))\}$  and:

$$\mathcal{E}^+(s, p) = \begin{cases} \mathcal{E}(s, p) & \text{if } p \in (0, 1), \\ \mathcal{E}(s, 0) + \mathbb{R}_+ \times \{0\} & \text{if } p = 0, \\ \mathcal{E}(s, 1) + \{0\} \times \mathbb{R}_+ & \text{if } p = 1. \end{cases}$$

Observe that:

- $\text{co}_v(G_s) = G_s$  for every  $s$ , which implies that the equilibrium payoffs are the same with or without cheap talk;
- Since  $|T| = 2$ , it is without loss of generality in terms of sender equilibrium payoffs to consider equilibria in which at most two signals are used (and, by the observation above, only one cheap talk message). So the equilibrium is either pooling on some signal  $\underline{s}$ , or it uses two signals  $\underline{s}$  and  $\bar{s}$  with  $\bar{s} > \underline{s}$ .

Figure 2 represents the projections of the graphs of  $\mathcal{E}^+(s, \cdot)$  on the sender's vector payoffs for three possible signals,  $s = 0$ ,  $s = \underline{s} > 0$  and  $s = \bar{s} > \underline{s}$ . The points in the positive orthant (i.e., satisfying the INTIR condition, represented by the gray area) in which the graphs cross each other are represented by orange dots and are candidates for incentive-compatible splittings.

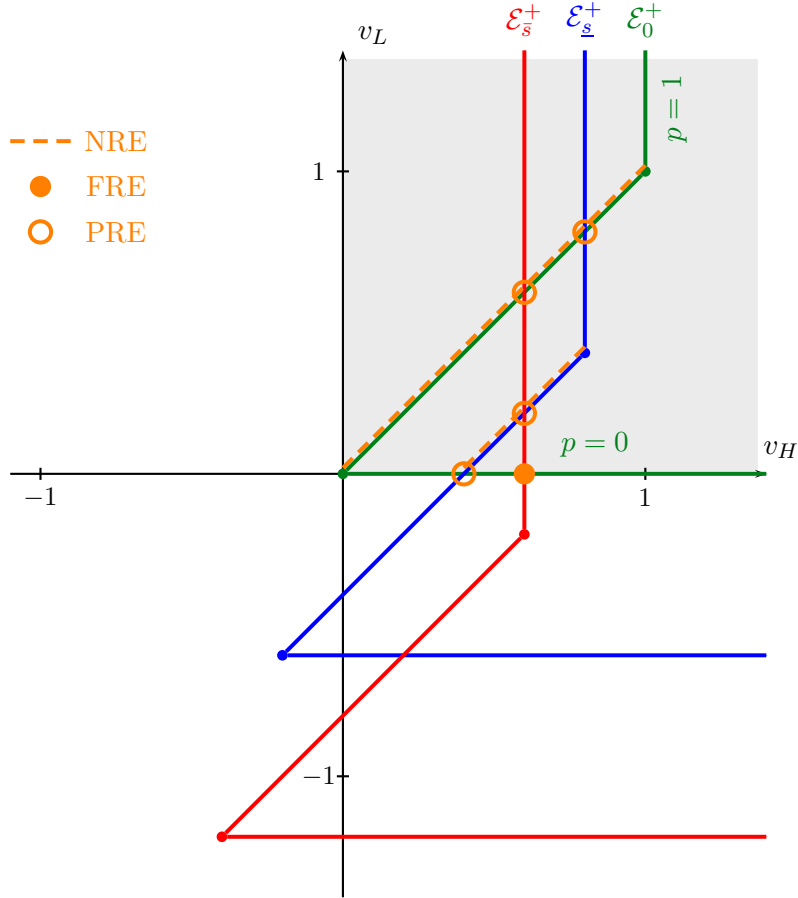


Figure 2: Graphs of  $\mathcal{E}^+(s, \cdot)$  (projected on the space of the sender vector payoffs, denoted  $\mathcal{E}_s^+$ ) in Spence job market signaling for  $s = 0$  (in green),  $s = \underline{s}$  (in blue) and  $s = \bar{s}$  (in red).

**Non-revealing equilibria** Vector  $v$  is a non-revealing (pooling) equilibrium payoff iff  $v$  is positive (the INTIR condition) and  $(v, p_0) \in \mathcal{E}_s(p_0)$  for some  $s$ , in which case signal  $s$  is sent with probability one by every sender's type. Hence, there is a non-revealing equilibrium on signal  $s$  if and only if  $f(p_0) - c(s, t_L) \geq 0$ , which can be written as  $s \leq \bar{s}^{NR}(p_0)$ , where  $\bar{s}^{NR}(p_0) > 0$  is strictly increasing in  $p_0$  and is the unique solution of  $f(p_0) - c(\bar{s}^{NR}(p_0), t_L) = 0$ . In Figure 2, the non-revealing equilibrium payoff candidates for some prior  $p_0$  are represented by orange dashed lines. On the figure, we see that: there is no non-revealing equilibrium with signal  $s = \bar{s}$  regardless of the prior; there is a range of high enough priors with a non-revealing equilibrium with signal  $\underline{s}$ ; and there is a non-revealing equilibrium with signal  $s = 0$  for all priors.

**Fully-revealing equilibria** Vector  $v$  is a fully-revealing equilibrium payoff iff  $v \in \mathbb{R}_+^2$ ,  $(v, 0) \in \mathcal{E}^+(s_L, 0)$  and  $(v, 1) \in \mathcal{E}^+(s_H, 1)$  for some pair of signals  $(s_L, s_H)$ . Such an equilibrium corresponds to a splitting of  $p_0$  on  $p = 0$  with signal  $s_L$  and  $p = 1$  with signal  $s_H$ . In Figure 2, such a fully revealing payoff

vector is a point where an horizontal line (corresponding to  $\mathcal{E}_{s_L}^+$ ) crosses a vertical line (corresponding to  $\mathcal{E}_{s_H}^+$ ). Such a point exists only for  $s_H > s_L$ , and to satisfy the INTIR condition we must have  $s_L = 0$  (because for other signals the horizontal line is not in the positive orthant). In the figure, there is a fully revealing equilibrium with  $s_H = \bar{s}$  (represented by the filled orange dot) but not with  $s_H = \underline{s}$ . More generally, such a fully revealing equilibrium exists if and only if:

$$\mathcal{E}(s_H, 1) \subseteq \mathbb{R}_+ \times \mathbb{R}_-,$$

that is:

$$f(1) - c(s_H, t_H) \geq 0 \text{ and } f(1) - c(s_H, t_L) \leq 0,$$

which we can rewrite:  $c(s_H, t_H) \leq f(1) \leq c(s_H, t_L)$ . The best fully revealing equilibrium for the sender is obtained with signal  $s_H = s_H^*$  where  $c(s_H^*, t_L) = f(1)$ . The worst fully revealing equilibrium for the sender is obtained with signal  $s_H = s_H^\#$  where  $c(s_H^\#, t_H) = f(1)$ . Hence, a fully revealing equilibrium always exists, the set of fully revealing equilibria is characterized by the splittings on  $p = 0$  with signal  $s = 0$  and  $p = 1$  with signal  $s = s_H$  for every  $s_H \in [s_H^*, s_H^\#]$ , and  $v$  is a fully revealing equilibrium payoff vector if and only if  $0 \leq v_H \leq f(1) - c(s_H^*, t_H)$  and  $v_L = 0$ .

**Partially-revealing equilibria** Vector  $v$  is a partially-revealing equilibrium payoff iff  $v \in \mathbb{R}_+^2$ ,  $(v, p') \in \mathcal{E}^+(s', p')$  and  $(v, p'') \in \mathcal{E}^+(s'', p'')$  for some pair of signals  $(s', s'')$  and some pair of beliefs  $(p', p'')$  with  $p' < p_0 < p''$  and either  $p' > 0$  or  $p'' < 1$  (if  $p' = 0$  and  $p'' = 1$  we get a fully revealing equilibrium as characterized in the previous paragraph). Such equilibrium candidates are represented in Figure 2 by the orange circles. As can be observed in the figure, there is no such equilibrium with  $0 < p' < p'' < 1$  because every such point is either on an horizontal line (with belief  $p = 0$ ) or a vertical line (with belief  $p = 1$ ). This means that one sender type randomizes on two signals and the other type does not randomize. Contrary to what we observed in a fully revealing equilibrium, it is possible that the two signals sent with strictly positive probability are both strictly positive. In the figure, this corresponds to the point where the red curve crosses the blue curve in the positive orthant, with payoff vector  $v = (0.6, 0.2)$ . Such an equilibrium exists for all high enough priors, and corresponds to a splitting on  $p' \in (0, p_0)$  with signal  $s' = \underline{s}$  and on  $p'' = 1$  with signal  $s'' = \bar{s}$ , meaning that type  $t_H$  randomizes between signal  $\bar{s}$  and  $\underline{s}$  and type  $t_L$  sends signal  $\underline{s}$  with probability 1. Two other partially revealing equilibria with similar properties appear in the figure with signals  $(s', s'') = (0, \underline{s})$  and  $(s', s'') = (0, \bar{s})$  and they exist for all high enough priors. An alternative type of partially revealing equilibrium is represented by the crossing point of the green and blue curve with  $v_L = 0$ . Such an equilibrium exists for all low enough priors, and corresponds to a splitting on  $p' = 0$  with signal  $s' = 0$  and on  $p'' \in (p_0, 1)$  with signal  $s'' = \underline{s}$ , meaning that type  $t_H$  sends signal  $\underline{s}$  with probability 1 and type  $t_L$  randomizes between signal  $\underline{s}$  and 0.

**Sender preferred equilibria** It is clear from the analysis above that the best non-revealing equilibrium for the sender is the one using signal  $s = 0$ , i.e., the worker pools on no education. Such an equilibrium also dominates (at least weakly) any partially revealing equilibrium for both sender types, and gives payoff  $v_L^{NR} = v_H^{NR} = f(p_0)$ . The best fully revealing equilibrium gives payoff  $v_L^{FR} = 0$  to type  $t_L$  and payoff  $v_H^{FR} = f(1) - c(s_H^*, t_H)$  to type  $t_H$ . Letting  $p^*$  be the unique prior satisfying  $f(p^*) = f(1) - c(s_H^*, t_H)$ , we get that when  $p_0 > p^*$  both sender types prefer the best non-revealing equilibrium, and when  $p_0 < p^*$  there is a conflict between type  $t_L$  and  $t_H$  because type  $t_L$  prefers the pooling equilibrium and type  $t_H$  prefers the fully revealing equilibrium.

## 5 Further comments and results

### 5.1 Comparison with commitment

If the sender can commit ex-ante to his messaging strategy, then the incentive compatibility conditions for the sender, as well as the interim individual rationality condition, are no longer relevant, and only the sequential rationality of the receiver matters. Thus, an interim payoff vector  $v$  is feasible in the signaling game with commitment if and only if  $(v, p_0) \in \text{co}(G)$ , and the maximum ex-ante expected utility of the sender is given by the concave closure  $\text{cav} \max_{s \in S} w(s, p)$ , where  $w(s, p) = \max_{y \in Y(s, p)} \sum_t p(t) u(y, s, t)$  for every  $s$ , and  $\text{cav}$  is the concavification operator, i.e.,  $\text{cav} f$  is the lowest (pointwise) concave function above  $f$ . When signals are payoff-irrelevant, or when  $|S| = 1$ , we directly obtain the characterization of Kamenica and Gentzkow (2011).

In the signaling game with commitment and without cheap talk messages, an interim payoff vector  $v$  is feasible if and only if  $(v, p_0)$  belongs to the *join* of the family of sets  $\{G_s : s \in S\}$  which are convex combinations  $(v, p_0) = \sum_s \lambda_s (v_s, p_s)$  where each point  $(v_s, p_s)$  belongs to  $G_s$ . This latter setting is analyzed by Boleslavsky and Shadmehr (2023).

### 5.2 Transparent motives

In this section, we consider a signaling game with *transparent motives* whereby the sender's utility is type independent, and we simply denote it by  $u : A \times S \rightarrow \mathbb{R}$ . The interim equilibrium payoff of the sender is then the same regardless of his type and we denote it  $v \in \mathbb{R}$ . With some abuse of notation, the set  $G$  can be written as:

$$G = \{(v, p) \in \mathbb{R} \times \Delta(T) : \exists s \in S, y \in Y(s, p) \text{ s.t. } v = u(y, s)\}.$$

Then, our conditions reformulate as follows:

- $(v, p_0) \in \text{co}_v(G)$  if and only if for some  $K \in \{1, \dots, |T|\}$ , there exist  $(p_k)_{k=1}^K$  and  $(s_k)_{k=1}^K$ , with  $p_k \in \Delta(T)$  and  $s_k \in S$  for every  $k$ , such that  $p_0 \in \text{co}\{p_1, \dots, p_K\}$ ,  $y_k \in Y(s_k, p_k)$  and  $v = u(y_k, s_k)$  for every  $k$ ;
- $v \in \text{INTIR}$  if and only if  $v \geq \max_{s \in S} \min_{p \in \Delta(T)} \min_{y \in Y(s, p)} u(y, s)$ .

Hence, from Theorem 1,  $v$  is a PBE payoff of the signaling game if and only if:

$$v \geq \max_{s \in S} \min_{p \in \Delta(T)} \min_{y \in Y(s, p)} u(y, s),$$

and there exists a family  $(p_k, s_k, y_k)_{k=1}^K$  such that  $p_0 \in \text{co}\{(p_k)_{k=1}^K\}$ , and for all  $k$ ,  $y_k \in Y(s_k, p_k)$  and  $v = u(y_k, s_k)$ .

### Maximal equilibrium payoff

For every  $s \in S$  and  $p \in \Delta(T)$ , the best non-revealing payoff for the sender at  $p$  given signal  $s$  is denoted by:

$$w(s, p) := \max_{y \in Y(s, p)} u(y, s).$$

For every  $p \in \Delta(T)$ , let  $w(p) := \max_{s \in S} w(s, p)$  be the best non-revealing payoff at  $p$  for the sender. For every function  $f : \Delta(T) \rightarrow \mathbb{R}$ , we denote by  $\text{qcav} f$  the smallest (pointwise) quasi-concave function above  $f$  (the quasi-concavification of  $f$ ).



**Theorem 2.** *With transparent motives, the maximal PBE payoff of the sender in the signaling game is  $\text{qcav } w(p_0)$ .*

If we further assume that signals are payoff-irrelevant, then the signaling game is a cheap talk game with transparent motives as in Lipnowski and Ravid (2020) and the characterization of Theorem 2 coincides with their Theorem 2. The proof of Theorem 2 can be found in Appendix A.

### Signaling without cheap talk

A signaling game is called *without cheap talk* if  $|M| = 1$ .

**Theorem 3.** *Under transparent motives, if for each  $s$ ,  $p \mapsto w(s, p)$  is quasi-concave, then every PBE payoff  $v \geq w(p_0)$  of the signaling game is also a PBE payoff of the signaling game without cheap talk.*

The proof of Theorem 3 can be found in Appendix B.

### 5.3 Zero-sum signaling games

In this section, we consider the particular case of zero-sum signaling games where  $u_R(a, s, t) = -u(a, s, t)$  for all  $(a, s, t) \in A \times S \times T$ . Then:

$$Y(s, p) = \arg \min_{y \in \Delta(A)} u(y, s, p).$$

Let  $\mathbf{w}(p) := \max_{s \in S} \min_{y \in \Delta(A)} u(y, s, p)$  be the maxmin value of the zero-sum signaling game where the sender does not know the type and  $p$  is the common prior. The concave closure of  $\mathbf{w}(\cdot)$  is:

$$\text{cav } \mathbf{w}(p) = \max \left\{ \sum_k \lambda_k \mathbf{w}(p_k) : \lambda_k \geq 0, p_k \in \Delta(T), \sum_k \lambda_k = 1, \sum_k \lambda_k p_k = p \right\}.$$

We say that a splitting  $(\lambda_k, p_k)_k$  of  $p_0$  is optimal if it achieves  $\text{cav } \mathbf{w}(p_0)$  in this maximisation problem.

**Theorem 4.** *Assume that the game is zero-sum.*

1. *The vector  $v \in \mathbb{R}^T$  is a PBE payoff if and only if:*

$$\sum_t p_0(t) v_t = \text{cav } \mathbf{w}(p_0) \text{ and } \forall p \in \Delta(T), \sum_t p(t) v_t \geq \mathbf{w}(p). \quad (14)$$

*Condition (14) holds if and only if  $v$  belongs to the super-differential of  $\text{cav } \mathbf{w}(\cdot)$  at  $p_0$ . Thus, the equilibrium payoff vector is unique if and only if  $\text{cav } \mathbf{w}$  is differentiable at  $p_0$ .*

*In each equilibrium: the sender chooses an optimal splitting  $(\lambda_k, p_k)_k$  of  $p_0$ , and for each  $k$ , a signal  $s_k$  that achieves the maximum in  $\mathbf{w}(p_k)$ ; for each signal  $s$ , the receiver chooses a mixed action  $y_s$  such that for all  $t$ ,  $u(y_s, s, t) \leq v_t$ .*

2. *Any PBE payoff vector is also a PBE payoff in the signaling game without cheap talk.*

The first point is a generalization of the results of Ponsard and Zamir (1973). The proof of Theorem 4 can be found in Appendix C.

### Example: simple poker

Consider a version of the poker example of Section 4.1 where the sender has only two signals  $s = 0$  or  $s = 2$ . Players' payoffs as a function of the state  $t \in T = \{t_L, t_H\}$ , the signal  $s \in S = \{0, 2\}$ , and the receiver's action  $a \in \{f, c\}$  are given by the following tables:

$s = 0$		$s = 2$	$f$	$c$
$t_H$	$-1, 1$	$t_H$	$1, -1$	$3, -3$
$t_L$	$-1, 1$	$t_L$	$1, -1$	$-3, 3$

We have

$$Y(2, p) = \begin{cases} \{c\} & \text{if } p < \frac{2}{3}, \\ \Delta(\{c, f\}) & \text{if } p = \frac{2}{3}, \\ \{f\} & \text{if } p > \frac{2}{3}. \end{cases}$$

It follows that:

$$\mathbf{w}(p) = \max \left\{ -1, \min \{1, 3p - 3(1 - p)\} \right\}.$$

As seen on Figure 3, the concave closure  $\text{cav } \mathbf{w}$  is differentiable at all points but  $p = \frac{2}{3}$ . Thus for  $p < \frac{2}{3}$ , the unique equilibrium payoff vector is  $(v_H, v_L) = (2, -1)$ , for  $p > \frac{2}{3}$ , the unique equilibrium payoff vector is  $(v_H, v_L) = (1, 1)$ . For  $p = \frac{2}{3}$  any vector  $(\frac{3-v_L}{2}, v_L)$  for  $v_L \in [-1, 1]$  is an equilibrium payoff.

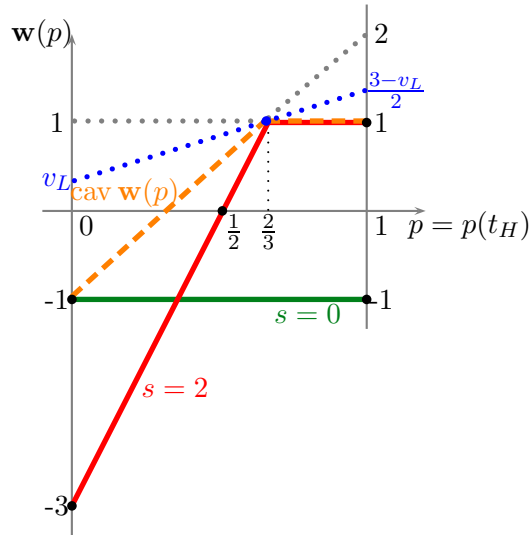


Figure 3: Simplified poker with  $s = 0, 2$ .

## References

- AUMANN, R. J. AND S. HART (2003): “Long Cheap Talk,” *Econometrica*, 71, 1619–1660.
- AUMANN, R. J. AND M. MASCHLER (1966): “Game Theoretic Aspects of Gradual Disarmament,” Report of the U.S. Arms Control and Disarmament Agency, ST-80, Chapter V, pp. 1–55.
- AUMANN, R. J. AND M. B. MASCHLER (1995): *Repeated Games of Incomplete Information*, Cambridge, Massachusetts: MIT Press.
- BOLESZLAVSKY, R. AND M. SHADMEHR (2023): “Signaling With Commitment,” *arXiv preprint arXiv:2305.00777*.
- CHAKRABORTY, A. AND R. HARBAUGH (2010): “Persuasion by Cheap Talk,” *American Economic Review*, 100, 2361–2382.
- FORGES, F. (1984): “Note on Nash equilibria in repeated games with incomplete information,” *International Journal of Game Theory*, 13, 179–187.
- (1990): “Equilibria with Communication in a Job Market Example,” *Quarterly Journal of Economics*, 105, 375–398.
- (1994): “Non-Zero Sum Repeated Games and Information Transmission,” in *Essays in Game Theory: In Honor of Michael Maschler*, ed. by N. Megiddo, Springer-Verlag.
- (2020): “Games with incomplete information: from repetition to cheap talk and persuasion,” *Annals of Economics and Statistics*, 3–30.
- FORGES, F. AND F. KOESSLER (2005): “Communication Equilibria with Partially Verifiable Types,” *Journal of Mathematical Economics*, 41, 793–811.
- GRAFEN, A. (1990): “Biological Signals as Handicaps,” *Journal of Theoretical Biology*, 144, 517–546.
- HART, S. (1985): “Nonzero-Sum Two-Person Repeated Games with Incomplete Information,” *Mathematics of Operations Research*, 10, 117–153.
- KAMENICA, E. AND M. GENTZKOW (2011): “Bayesian Persuasion,” *The American Economic Review*, 101, 2590–2615.
- KREPS, D. M. AND J. SOBEL (1994): “Signalling,” in *Handbook of Game Theory*, ed. by R. J. Aumann and S. Hart, Elsevier Science B. V., vol. 2, chap. 25, 849–867.
- KREPS, D. M. AND R. WILSON (1982): “Sequential Equilibria,” *Econometrica*, 50, 863–894.
- LIPNOWSKI, E. AND D. RAVID (2020): “Cheap talk with transparent motives,” *Econometrica*, 88, 1631–1660.
- MANELLI, A. M. (1996): “Cheap talk and sequential equilibria in signaling games,” *Econometrica: Journal of the Econometric Society*, 917–942.
- MAS-COLELL, A., M. D. WHINSTON, AND J. R. GREEN (1995): *Microeconomic Theory*, New York: Oxford University Press.
- PONSSARD, J. P. AND S. ZAMIR (1973): “Zero-sum sequential games with incomplete information,” *International Journal of Game Theory*, 2, 99–107.

- SION, M. (1958): “On general minimax theorems,” *Pacific Journal of Mathematics*, 1, 171–176.
- SOBEL, J. (2020): “Signaling games,” *Complex Social and Behavioral Systems: Game Theory and Agent-Based Models*, 251–268.
- SPENCE, A. M. (1973): “Job Market Signaling,” *Quarterly Journal of Economics*, 87, 355–374.
- ZAHAVI, A. (1975): “Mate selection – A selection for a handicap,” *Journal of Theoretical Biology*, 53.

## A Proof of Theorem 2

To prove Theorem 2, we first establish the next proposition, which is an analogue of Theorem 1 in Lipnowski and Ravid (2020) for signaling games with transparent motives.

**Proposition 1.** *Let  $(p_k)_{k=1}^K$  be a set of posteriors with  $p_k \in \Delta(T)$  for every  $k = 1, \dots, K$  such that  $p_0 \in \text{co}\{p_1, \dots, p_K\}$  and  $\min_k w(p_k) \geq w(p_0)$ . Then,  $\min_k w(p_k)$  is a PBE payoff of the signaling game with transparent motives.*

*Proof of Proposition 1.* Suppose that  $p_0 \in \text{co}\{p_1, \dots, p_K\}$ . First, we argue that, for every  $j = 1, \dots, K$  and every  $\alpha \in [0, 1]$ :

$$p_0 \in \text{co}\{(1 - \alpha)p_0 + \alpha p_j, p_k, k \neq j\}.$$

To see this, note that:

$$p_0 = \sum_k \lambda_k p_k = \lambda_j p_j + (1 - \lambda_j) \sum_{k \neq j} \frac{\lambda_k}{1 - \lambda_j} p_k := \lambda_j p_j + (1 - \lambda_j) \bar{p}_j.$$

Consider now  $q = (1 - \alpha)p_0 + \alpha p_j$  and assume  $0 < \alpha < 1$  (the cases  $\alpha = 0$  or  $1$  are clear). It is easy to check that:

$$p_0 = \frac{\lambda_j}{\lambda_j + \alpha(1 - \lambda_j)} q + \frac{\alpha(1 - \lambda_j)}{\lambda_j + \alpha(1 - \lambda_j)} \bar{p}_j.$$

Now, fix  $(p_k)_k$  such that  $p_0 \in \text{co}\{p_1, \dots, p_K\}$  and  $w^* = \min_k w(p_k) \geq w(p_0)$ . Suppose that for some  $j$ ,  $w(p_j) > w^*$ .

**Lemma 2.** *There exists  $\alpha_j \in [0, 1]$ ,  $s_j \in S$  and  $y_j \in Y(s_j, (1 - \alpha_j)p_0 + \alpha_j p_j)$  such that*

$$u(s_j, (1 - \alpha_j)p_0 + \alpha_j p_j) = w^*.$$

*Proof of Lemma 2.* For  $q \in \Delta(T)$ , denote  $\underline{v}(q) = \max_{s \in S} \min_{y \in Y(s, q)} u(y, s)$  and

$$W(q) = \{u(y, s) : s \in S, y \in Y(s, q), u(y, s) \geq \underline{v}(q)\}.$$

**Claim 1.** *The correspondence  $W$  is Kakutani (i.e., u.h.c. with non-empty, convex, compact values).*

*Proof of Claim 1.*  $W(q)$  is non-empty because  $w(q) \in W(q)$ . Verification of u.h.c. is routine: Take  $q^n \rightarrow q$ ,  $s^n \rightarrow s$ ,  $y^n \in Y(s^n, q^n)$ ,  $y^n \rightarrow y$  and for each  $s'$ , there exists  $y'^n \in Y(s', q^n)$  such that  $u(y^n, s^n) \geq u(y'^n, s')$ . Since  $u(\cdot, \cdot)$  is continuous, from the maximum theorem,  $Y(\cdot, \cdot)$  is u.h.c. and thus  $y \in Y(s, q)$ . For each  $s'$ , take a converging subsequence to make sure that  $y'^n \rightarrow y'$  and then  $y' \in Y(s', q)$ . So for each  $s'$ , there exists  $y' \in Y(s', q)$  such that  $u(y, s) \geq u(y', s')$ , i.e.,  $u(y, s) \geq \underline{v}(q)$ .

We prove now that  $W(q)$  is convex. Observe that  $W(q) = \cup_{s \in S} W_s(q)$  with:

$$W_s(q) = \{u(y, s) : y \in Y(s, q), u(y, s) \geq \underline{v}(q)\}.$$

More precisely,  $W(q) = \cup_{s \in S, W_s(q) \neq \emptyset} W_s(q)$ . For each  $s$ ,  $Y(s, q)$  is a convex set of mixed actions and  $y \mapsto u(y, s)$  is linear on this set, thus  $W_s(q)$  is a compact interval of  $\mathbb{R}$ . To prove that the union of those intervals is convex, it is enough to show that any pair of them has a non-empty intersection. Thus, take  $s_1, s_2$  such that  $W_{s_1}(q)$  and  $W_{s_2}(q)$  are non-empty and suppose by contradiction that  $W_{s_1}(q) \cap W_{s_2}(q) = \emptyset$ . Then, one of those intervals, say  $W_{s_2}(q)$ , is “above” the other, that is:

$$\max W_{s_1}(q) < \min W_{s_2}(q);$$

or equivalently:

$$\max\{u(y, s_1) : y \in Y(s_1, q), u(y, s_1) \geq \underline{v}(q)\} < \min\{u(y, s_2) : y \in Y(s_2, q), u(y, s_2) \geq \underline{v}(q)\}.$$

Consider then the interval:

$$W'_{s_2}(q) = \{u(y, s_2) : y \in Y(s_2, q)\}.$$

We have  $W_{s_2}(q) \subseteq W'_{s_2}(q)$  and  $\min W'_{s_2}(q) \leq \underline{v}(q)$ . Then, either  $\max W_{s_1}(q) < \min W'_{s_2}(q)$ , thus  $\max W_{s_1}(q) < \underline{v}(q)$  which contradicts that  $W_{s_1}(q)$  is non-empty. Or  $\max W_{s_1}(q) \geq \min W'_{s_2}(q)$ , but then  $W_{s_1}(q) \cap W'_{s_2}(q) \neq \emptyset$ . This is also a contradiction because  $W_{s_1}(q) \cap W'_{s_2}(q) \subseteq W_{s_1}(q) \cap W_{s_2}(q)$ . This concludes the proof of Claim 1.  $\square$

Now, it follows that the correspondence  $\alpha \mapsto W((1 - \alpha)p_0 + \alpha p_j)$  from  $[0, 1]$  to  $\mathbb{R}$  is also Kakutani. From Lemma 3 in Lipnowski and Ravid, its image is an interval. This proves Lemma 2 because  $w(p_j) \in W(p_j)$ ,  $w(p_0) \in W(p_0)$  and  $w^* \in [w(p_0), w(p_j)]$ .  $\square$

We conclude the proof of Proposition 1. For each  $j$  such that  $w(p_j) > w^*$ , we replace  $p_j$  by  $(1 - \alpha_j)p_0 + \alpha_j p_j$  given by Lemma 2. Together with the family of  $s_j, y_j$ , this defines an equilibrium with payoff  $w^*$ .  $\square$

Consider now the maximal equilibrium payoff  $u^*(p_0)$  of the sender at  $p_0$ . Observe that:

$$u^*(p_0) = \max \left\{ v \in \mathbb{R} : \exists (p_k)_k, p_0 \in \text{co} \{(p_k)_k\} \text{ and } \forall k, \exists s_k \in S, \exists y_k \in Y(s_k, p_k), v = u(y_k, s_k) \right\}.$$

The right-hand-side of this equality is greater or equal than  $w(p_0)$ , by considering the non-revealing splitting ( $\forall k, p_k = p_0$ ) and  $w(p_0) \geq \max_{s \in S} \min_{p, y \in Y(s, p)} u(y, s)$ . Thus the INTIR condition is superfluous in defining  $u^*(p_0)$ . Also,  $u^*(p_0) \geq w(p_0)$ .

We now derive two consequences of Proposition 1.

**Corollary 5.** *For any family  $(p_k)_k$  such that  $p_0 \in \text{co} \{(p_k)_k\}$ , we have  $u^*(p_0) \geq \min_k w(p_k)$ .*

*Proof of Corollary 5.* If  $\min_k w(p_k) \leq w(p_0)$ , this follows from  $u^*(p_0) \geq w(p_0)$ . If  $\min_k w(p_k) > w(p_0)$ , from Proposition 1,  $\min_k w(p_k)$  is an equilibrium payoff, thus less or equal to  $u^*(p_0)$ .  $\square$

**Corollary 6.** *Let  $(p_k^*, s_k^*, y_k^*)_k$  be an optimal family in the maximisation problem defining  $u^*(p_0)$ . Then  $u^*(p_0) = \min_k w(p_k^*)$ .*

*Proof of Corollary 6.* For each  $k$ ,  $u^*(p_0) = u(y_k^*, s_k^*) \leq w(p_k^*)$ , thus  $u^*(p_0) \leq \min_k w(p_k^*)$ . On the other hand,  $\min_k w(p_k^*)$  is an equilibrium payoff from Proposition 1, thus it is less or equal to  $u^*(p_0)$ .  $\square$

The previous corollaries lead to the following proposition.

**Proposition 2.** *The function  $p_0 \mapsto u^*(p_0)$  is quasi-concave.*

*Proof of Proposition 2.* We prove now that for any family  $(p_k)_k$  such that  $p_0 \in \text{co} \{(p_k)_k\}$ ,  $u^*(p_0) \geq \min_k u^*(p_k)$ . For each  $k$ , let  $(p_{kj}^*, s_{kj}^*, y_{kj}^*)_j$  be an optimal family in the maximisation problem defining  $u^*(p_k)$ . Since  $p_0 \in \text{co} \{(p_k)_k\}$  and  $p_k \in \text{co} \{(p_{kj}^*)_j\}$ , then  $p_0 \in \text{co} \{(p_{kj}^*)_{k,j}\}$ . Then from Corollary 5,

$$u^*(p_0) \geq \min_k \min_j w(p_{kj}^*)$$

and from Corollary 6,  $\min_j w(p_{kj}^*) = u^*(p_k)$ . Thus,  $u^*$  is quasi-concave.  $\square$

*Proof of Theorem 2.* We are now in position to prove the theorem. We know that the maximal equilibrium payoff  $u^*(p_0)$  of the sender is a quasi-concave function of  $p_0$  which is greater than or equal to  $w(p_0)$ . We argue that this is the smallest such quasi-concave function: if  $f$  is quasi-concave and  $f(p_0) \geq w(p_0)$  for all  $p_0$ , then  $f(p_0) \geq u^*(p_0)$  for all  $p_0$ . To see this, let  $(p_k^*, s_k^*, y_k^*)_k$  be an optimal family in the maximisation problem defining  $u^*(p_0)$ . From Corollary 6,  $u^*(p_0) = \min_k w(p_k^*)$ . Then,

$$u^*(p_0) = \min_k w(p_k^*) \leq \min_k f(p_k^*) \leq f(p_0),$$

since  $f$  is quasi-concave. This concludes the proof.  $\square$

## B Proof of Theorem 3

*Proof.* Let  $v \geq w(p_0)$  be a PBE payoff of the signaling game. We construct a PBE with payoff  $v$  such that  $k, k', p_k \neq p_{k'}$  and  $s_k \neq s_{k'}$ . In words, all the information transmitted by the equilibrium is contained in the signal: if two signals are the same, then the posterior beliefs are also the same.

Since  $v$  is a PBE payoff of the signaling game, there exists a finite family  $(p_k, s_k, y_k)_{k=1}^K$  such that  $p_0 \in \text{co} \{(p_k)_{k=1}^K\}$  and for all  $k$ ,  $y_k \in Y(s_k, p_k)$  and  $v = u(y_k, s_k)$ .

Since the statement is clear for  $v = w(p_0)$ , take  $v > w(p_0)$ . For each  $s \in S' = \{s_k : k = 1, \dots, K\}$ , denote  $K_s = \{k : s_k = s\}$ . For some convex combination coefficients  $\lambda_k > 0$ , we have:

$$p_0 = \sum_k \lambda_k p_k = \sum_{s \in S'} \lambda_s \sum_{k \in K_s} \frac{\lambda_k}{\lambda_s} p_k = \sum_s \lambda_s p_s,$$

with  $\lambda_s = \sum_{k \in K_s} \lambda_k$  and  $p_s = \sum_{k \in K_s} \frac{\lambda_k}{\lambda_s} p_k$ , so  $p_0 \in \text{co} \{(p_s)_{s \in S'}\}$ .

For each signal  $s$ , denote:

$$D_s = \{p \in \Delta(T) : \exists y \in Y(s, p), u(y, s) \geq v\} = \{p \in \Delta(T) : w_s(p) \geq v\}.$$

Since  $w_s$  is quasi-concave,  $D_s$  is convex, it is thus a non-empty convex and compact set (since  $w_s$  is u.s.c). From the equilibrium condition, for all  $s \in S'$  and all  $k \in K_s$ ,  $u(y_k, s_k) = v$ , thus  $p_k \in D_s$ . Since  $D_s$  is convex, we also have  $p_s \in D_s$  for each  $s$ .

Thus for each  $s$  we have:

$$w_s(p_0) \leq w(p_0) < v \leq w_s(p_s).$$

Consider then for  $\alpha \in [0, 1]$ ,

$$F_s(\alpha) = \{u(y, s) : y \in Y(s, (1 - \alpha)p_0 + \alpha p_s)\}.$$

Similarly to the proof of Lemma 2, this correspondance is Kakutani, thus its image is an interval. It follows that there exists  $\alpha_s \in [0, 1]$  and  $y_s \in Y(s, (1 - \alpha)p_0 + \alpha p_s)$  such that  $v = u(y_s, s)$ . We modify then the equilibrium by replacing all the  $(y_k, p_k)$  for  $k \in K_s$ , by  $(y_s, p_s)$ . This gives the desired conclusion.  $\blacksquare$

## C Proof of Theorem 4

*Proof.* (1) Let  $v \in \mathbb{R}^T$  be such that Condition (14) holds.

Consider an optimal splitting  $(\lambda_k, p_k)_k$  of  $p_0$  such that  $\text{cav } \mathbf{w}(p_0) = \sum_k \lambda_k \mathbf{w}(p_k)$  and let the sender choose a strategy that achieves this splitting. Precisely, there is a finite subset of signals and messages

$(s_k, m_k)_k$  and let  $\sigma : T \rightarrow \Delta(S \times M)$  be such that  $\sigma(s_k, m_k|t) = p_k(t)\lambda_k/p_0(t)$  and  $s_k$  achieves the maximum in  $\mathbf{w}(p_k)$ . For any strategy of the receiver playing action  $a_k$  after observing  $(s_k, m_k)$ , the subsequent payoff  $u(a_k, s_k, p_k)$  is greater or equal to  $\min_a u(a, s_k, p_k)$  and therefore it is greater or equal to  $\max_s \min_a u(a, s, p_k) = \mathbf{w}(p_k)$  from optimality of  $s_k$ . Thus, the ex-ante payoff is at least  $\sum_k \lambda_k \mathbf{w}(p_k) = \text{cav } \mathbf{w}(p_0)$ . This shows that the sender can guarantee  $\text{cav } \mathbf{w}(p_0)$  which is greater or equal to the maxmin value.

The inequalities of Condition (14) amount to:

$$\begin{aligned} 0 &\geq \max_p \max_s \min_y \left\{ u(y, s, p) - \sum_t p(t)v_t \right\} \\ &= \max_s \max_p \min_y \left\{ u(y, s, p) - \sum_t p(t)v_t \right\} \\ &= \max_s \min_y \max_p \left\{ u(y, s, p) - \sum_t p(t)v_t \right\}, \end{aligned}$$

where the second equality follows from the minmax Theorem (if  $A$  is not finite, see Sion (1958)). Therefore, for all  $s$ , there exists  $y_s$  such that for all  $p$ ,  $u(y_s, s, p) \leq \sum_t p(t)v_t$  and in particular for all  $t$ ,  $u(y_s, s, t) \leq v_t$ . Consider the strategy of the receiver that plays  $y_s$  after signal  $s$ . Since  $u(y_s, s, t) \leq v_t$  for all  $t$ , the ex-ante payoff is no more than  $\sum_t p_0(t)v_t = \text{cav } \mathbf{w}(p_0)$ , for any strategy of the sender. Thus,  $\text{cav } \mathbf{w}(p_0)$  is less or equal to the minmax value.

It follows that the minmax theorem holds and that the equilibria are given by the minmax and maxmin strategies described above.

Suppose now that  $v \in \mathbb{R}^T$  is an equilibrium payoff vector. Then from the arguments above, its ex-ante value  $\sum_t p_0(t)v_t$  has to be equal to  $\text{cav } \mathbf{w}(p_0)$  since a zero-sum game has a unique ex-ante equilibrium payoff. Now,  $v$  must belong to INTIR, so we have:

$$\forall s, \exists y, \forall t, v_t \geq u(y, s, t).$$

Therefore, for all  $s$ , there exists  $y$  such that for all  $p$ ,  $\sum_t p(t)v_t \geq u(y, s, p)$ . It follows that for each  $p$ ,  $\sum_t p(t)v_t \geq \max_s \min_y u(y, s, p) = \mathbf{w}(p)$ . Hence Condition (14) is satisfied.

Notice that since  $\sum_t p(t)v_t$  is a linear function of  $p$ , the inequalities  $\forall p \in \Delta(T), \sum_t p(t)v_t \geq \mathbf{w}(p)$  are equivalent to  $\forall p \in \Delta(T), \sum_t p(t)v_t \geq \text{cav } \mathbf{w}(p)$ . Condition (14) is thus equivalent to  $v$  belonging to the super-differential of  $\text{cav } \mathbf{w}$  at  $p_0$ .

(2) Next, we prove that cheap talk messages are not necessary. Take an equilibrium of the game given by a family  $(\lambda_k, p_k, s_k, y_k)_k$  with cardinality  $K$ , and with payoff vector  $v = (v_t)_t$ , and suppose that there are two indices  $j \neq \ell$  such that  $s_j = s_\ell := \bar{s}$ ,  $\lambda_j > 0$ ,  $\lambda_\ell > 0$ . We will merge  $j$  and  $\ell$  together replacing  $p_j$  and  $p_\ell$  by  $\bar{p} = \frac{\lambda_j}{\lambda_j + \lambda_\ell} p_j + \frac{\lambda_\ell}{\lambda_j + \lambda_\ell} p_\ell$ . We will choose  $\bar{y}$  in such a way that the family:

$$\{(\lambda_k, p_k, s_k, y_k)_{k \neq j, \ell}\} \cup \{(\lambda_j + \lambda_\ell, \bar{p}, \bar{s}, \bar{y})\},$$

is an equilibrium with the same payoff vector  $v$ .

We start by observing that  $\mathbf{w}(p)$  must be linear on the segment  $[p_j, p_\ell]$ , that is:

$$\forall \mu \in [0, 1], \mathbf{w}(\mu p_j + (1 - \mu)p_\ell) = \mu \mathbf{w}(p_j) + (1 - \mu) \mathbf{w}(p_\ell).$$

To see this, first assume that  $\mathbf{w}(\mu p_j + (1 - \mu)p_\ell) > \mu \mathbf{w}(p_j) + (1 - \mu) \mathbf{w}(p_\ell)$  for some  $\mu \in [0, 1]$ . This contradicts the optimality of the splitting because one could replace  $p_j$  and  $p_\ell$  by  $\mu p_j + (1 - \mu)p_\ell$  and



improve the value of  $\sum_k \lambda_k \mathbf{w}(p_k)$ . Thus,  $\mathbf{w}(\mu p_j + (1 - \mu)p_\ell) \leq \mu \mathbf{w}(p_j) + (1 - \mu)\mathbf{w}(p_\ell)$ . On the other hand:

$$\begin{aligned} \mathbf{w}(\mu p_j + (1 - \mu)p_\ell) &= \max_s \min_a u(a, s, \mu p_j + (1 - \mu)p_\ell) \\ &\geq \min_a u(a, \bar{s}, \mu p_j + (1 - \mu)p_\ell) \\ &\geq \mu \min_a u(a, \bar{s}, p_j) + (1 - \mu) \min_a u(a, \bar{s}, p_\ell) \\ &= \mu \mathbf{w}(p_j) + (1 - \mu)\mathbf{w}(p_\ell). \end{aligned}$$

The second inequality follows from concavity of the minimum of linear functions and the last equality from the optimality of  $\bar{s}$  at  $p_j$  and  $p_\ell$ . This implies that  $\bar{s}$  is optimal all along the segment  $[p_j, p_\ell]$  and thus  $\min_a u(a, s, \mu p_j + (1 - \mu)p_\ell)$  is an affine function of  $\mu$ . From this, we deduce that the set of optimal actions:

$$A_\mu^* := \arg \min_a u(a, s, \mu p_j + (1 - \mu)p_\ell),$$

is the same set  $A^*$  for all  $\mu \in (0, 1)$  and that all actions in  $A^*$  are also optimal at  $p_j$  and at  $p_\ell$  ( $\mu = 0, 1$ ).

Now, from the equilibrium conditions, for all  $t$  such that  $p_j(t) > 0$ , we have  $u(y_j, s, t) \geq u(y_\ell, s, t)$ , and thus  $u(y_j, s, p_j) \geq u(y_\ell, s, p_j)$ . Further, it must be the case that  $u(y_j, s, t) = u(y_\ell, s, t)$  for all  $t$  such that  $p_j(t) > 0$ . Otherwise, we would have that  $u(y_j, s, t) > u(y_\ell, s, t)$  for some  $t$  such that  $p_j(t) > 0$ , implying  $u(y_j, s, p_j) > u(y_\ell, s, p_j)$  and contradicting the optimality of  $y_j$  at  $p_j$ . We conclude that:

$$\forall t \text{ s.t. } p_j(t) > 0, u(y_j, s, t) = u(y_\ell, s, t) = u(a, s, t), \forall a \in A^*.$$

By a symmetrical argument:

$$\forall t \text{ s.t. } p_\ell(t) > 0, u(y_\ell, s, t) = u(y_j, s, t) = u(a, s, t), \forall a \in A^*.$$

The set of types  $t$  such that  $p_j(t) > 0$  or  $p_\ell(t) > 0$  is precisely the set of types that send signals  $s_j, s_\ell$  with positive probability. Since  $s_j = s_\ell$ , the receiver can play some  $\bar{y} \in A^*$  for both signals. This construction forms an equilibrium with a family of cardinality  $K - 1$  and the same vector of interim payoffs. Iterating the argument, we find that this payoff vector can be obtained with the property that if  $j \neq \ell$  then  $s_j \neq s_\ell$ . ■