Sequential Bayesian Persuasion*

Wenhao Wu†

Feburary, 2018

Abstract

This paper studies a Bayesian Persuasion model in which multiple senders sequentially persuade one Receiver. Players can always observe signaling rules of prior players and their realizations. I develop a recursive concavification method to characterize the set of SPEa. I prove the existence of a special type of equilibrium, called the Silent Equilibrium, where at most one sender designs a nontrivial signaling rule. Also, I show that in zero-sum games, the truth-telling information structure is always supported in equilibrium. Finally, I make comparisons with the simultaneous multi-sender Bayesian persuasion model of Gentzkow and Kamenica (2017b) and examine the impact of the order of persuasion. A geometric version of Blackwell's order is adopted for examination of informativeness.

Keywords: Bayesian Persuasion, Multiple Senders, Subgame Perfect Equilibrium, Markov Perfect Equilibrium, Communication, Infinite Game.

JEL Classification Codes: D82, D83

*I am indebeted to Andreas Blume for his encouragement and comments. I would also like to thank Martin Dufwenberg, Mark Walker, Asaf Plan, Peter Norman and participants at the U of Arizona theory lunch, the 28th International Conference on Game Theory at Stony Brook, and the Fall 2017 Midwest Economic Theory meeting, for their helpful comments. All errors are my own.

†University of Arizona. wuw@email.arizona.edu

1 Introduction

Bayesian persuasion, introduced by Kamenica and Gentzkow (2011), describes a situation in which a sender can fully commit to a signaling rule that maps states of nature into signals. When there is only one sender, he is able to manipulate the Receiver's belief and exploit the rent of persuasion. The Receiver follows the sender's suggestions, even though she realizes that the information is generated from a biased source.

When multiple senders get involved in persuasion, however, not only does the Receiver need to integrate information from different sources, but also the senders have to consider what information others disclose. Strategic interactions and conflicts of interest confound senders' behavior and require in-depth analysis. Gentzkow and Kamenica (2017b) study this class of Bayesian persuasion games with multiple senders who move simultaneously; in contrast, this paper is devoted to its counterpart in which senders move sequentially.

Assuming that previously designed signaling rules and their realizations are public, subsequent senders can correlate the probabilities of their signals with previous senders'.[1] Correlation makes this a coordinated information structure (Li and Norman 2017a, 2017b). From the perspective of informativeness, this setting implies that subsequent senders can select any weakly more transparent information structure than the existing one (Gentzkow and Kamenica 2017b).

I approach this problem by a recursive method that decomposes this multi-sender Bayesian persuasion game into separate single-sender games. Considering each sender's decision making, it is rational for him to update his belief based on previous information and evaluate the consequences of his persuasion as affected by reactions of subsequent players. Then, he confronts an updated "prior" and a current value referring to future equilibrium payoffs as if he is the only sender. Therefore, the equilibrium of the entire game is achieved by summarizing these decentralized analyses. Next, I present an example to illustrate the above idea.

Example 1: Judge, Prosecutor and Attorney

This example extends the judge-prosecutor example in Kamenica and Gentzkow (2011) by adding an attorney who defends the suspect. The game proceeds as follows: the prosecutor moves first [2] and picks a signaling rule that sends signal I or G randomly conditional on the underlying states *inno-*

[1] This is the same assumption on the space of signaling rules in Gentzkow and Kamenica (2017b).

[2] The equilibrium outcome is independent of the order in this example.

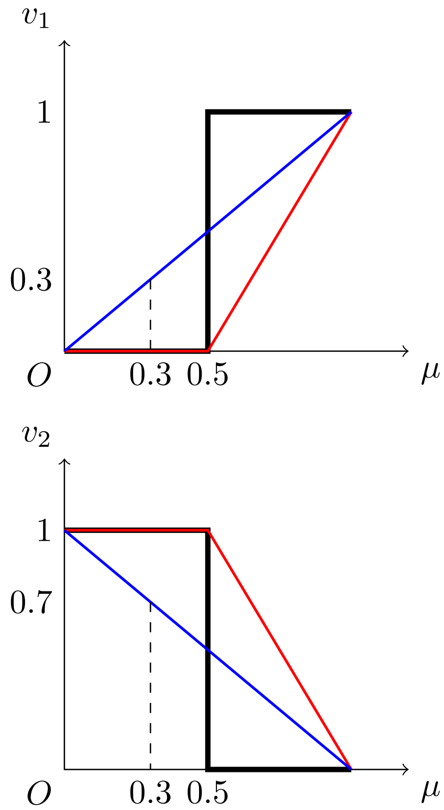


Figure 1: Judge-prosecutor-attorney persuasion game

cent or *guilty*. After observing the prosecutor's rule and signal, the attorney designs his own signaling rule that yields another signal. Finally, the judge reviews reports from both sides and makes a decision. No one knows the truth, while the common prior of the person's being *guilty* is 30% ($\mu_0 = 0.3$).

Suppose the judge wants to *convict* the *guilty* but *acquit* the *innocent*. She obtains 1 unit for doing so, 0 units if she fails. The prosecutor (attorney) only wants to have the person *convicted* (*acquitted*). The prosecutor's (attorney's) payoff is 1 for conviction (acquittal) and 0 otherwise. This is a Bayesian persuasion game with the state space $\{innocent, guilty\}$, the signal space $\{I, G\}$, and the action space $\{acquit, convict\}$. The belief held by players starts with μ_0 and keeps changing as the game proceeds. After the prosecutor's persuasion the belief is updated to μ_1 and the attorney's disclosure further leads to a new belief μ_2. I denote by $V_t^k(\mu_t)$ Sender k's value depending on the posterior of period t, where the superscript indexes the player and the subscript the period. A complete explanation of notation is in Section 3.1. Specifically, $\{V_0^k(\mu_0)\}_{\forall k}$ are equilibrium payoffs to the senders for a game starting with a prior μ_0.

Consider the final phase where the judge moves. Her optimal strategy for all possible beliefs, μ_2, she may hold at the end, can be described as: if $\mu_2 < 0.5$, she *acquits*; if $\mu_2 > 0.5$, she *convicts*; if $\mu_2 = 0.5$, she randomizes arbitrarily between *convict* and *acquit*. Endogenizing the judge's responses, the prosecutor's values can be defined as below:

$$V_2^1(\mu_2) = \begin{cases} 0 & \text{if } \mu_2 \in [0, 0.5), \\ [0, 1] & \text{if } \mu_2 = 0.5, \\ 1 & \text{if } \mu_2 \in (0.5, 1]. \end{cases}$$

Similarly, the attorney's payoff is defined as a function of final beliefs as below:

$$V_2^2(\mu_2) = \begin{cases} 1 & \text{if } \mu_2 \in [0, 0.5), \\ [0, 1] & \text{if } \mu_2 = 0.5, \\ 0 & \text{if } \mu_2 \in (0.5, 1]. \end{cases}$$

When it is the attorney's turn at the beginning of the second period, he has observed the prosecutor's persuasion and formed a new belief, μ_1. No matter which μ_1 is realized, the attorney acts as if he is the only sender in a Bayesian persuasion game with this new prior μ_1 and the value V_2^2. Hence, the attorney's value dependent on the result of the prosecutor's persuasion, V_1^2, is given by the concave closure of V_2^2 (the lowest concave function which is weakly higher than V_2^2), as the lower red curve shows. In equilibrium, the judge must break the tie by *acquitting* the person in favor of the attorney, otherwise, the attorney has a profitable deviation to a persuasion that randomizes between $\mu_2 = 1$ and $0.5 - \epsilon$, where ϵ is a sufficiently small positive number.[3]

$$V_1^2(\mu_1) = \begin{cases} 1 & \text{if } \mu_1 \in [0, \frac{1}{2}], \\ 2 - 2\mu_1 & \text{if } \mu_1 \in (\frac{1}{2}, 1]. \end{cases}$$

Furthermore, as the prosecutor and attorney play a zero sum game, the prosecutor's value defined on μ_1, V_1^1, can be obtained by subtracting V_1^2 from 1.

[3]Referring to Theorem 1, Section 4.

$$V_1^1(\mu_1) = \begin{cases} 0 & \text{if } \mu_1 \in [0, \frac{1}{2}], \\ -1 + 2\mu_1 & \text{if } \mu_1 \in (\frac{1}{2}, 1]. \end{cases}$$

With the knowledge of V_1^1, the prosecutor completely understands the consequence of his persuasion. Like the attorney, he is equivalently playing a single-sender Bayesian persuasion game with the prior μ_0 and value V_1^1. His initial value V_0^1 is indicated as the concave closure of V_1^1, shown as the blue line, which suggests that full revelation is his unique equilibrium strategy. Symmetrically, one can obtain the attorney's initial value, V_0^2, by substracting V_0^1 from 1.

$$V_0^1(\mu) = \mu, \mu \in [0, 1]$$

$$V_0^2(\mu) = 1 - \mu, \mu \in [0, 1]$$

Therefore, the equilibrium payoffs for both senders are $V_0^1(\mu_0) = 0.3$ and $V_0^2(\mu_0) = 0.7$. During this process, the equilibrium strategy is recovered by repeatedly imposing concave closures over relevant values. As a result, the prosecutor chooses a truth-telling signaling rule in equilibrium and the judge follows the prosecutor's suggestion. As one main contribution of this paper, I formally develop this method as the "recursive concavification method" in Section 4.

Besides the general technique, this paper solves for the Subgame Perfect Equilibrium which is more general than the equilibrium concepts in the existing literature, e.g. the Sender-preferred SPE and Markov Perfect Equilibrium (Kamenica and Gentzkow 2011; Gentzkow and Kamenica 2017b; Li and Norman 2017b; Ely 2017). Compared to those concepts, SPE has two advantages: 1) no behavioral restrictions and 2) exhausting the strategic interactions among players. As will be shown in Sections 4.5 and 5.3, the sets of equilibrium payoffs of the Sender-preferred SPE and MPE are proper subsets of that of SPE.

Notice that it is an equilibrium for the prosecutor to reveal the true state and the attorney to send a null signal. [4] In Section 6, I will show the pervasive existence of this type of equilibrium, called the Silent Equilibrium, where at most one sender reveals information. The existence of this equilibrium relies on the accessibility of an abundant set of signaling rules to the senders, which

[4] A signal from a babbling signaling rule that leaves the belief unchanged.

allows them to reveal beforehand what their followers would otherwise have revealed.

The economic implication of this example is also significant. In contrast with the previous judge-prosecutor game (Kamenica and Gentzkow, 2011), which incorporates only partial revelation in equilibrium, introducing a sender with an opposite interest to the previous sender improves information transparency significantly. More strikingly, this result can be generalized to any pair of senders who have zero sum utilities (Section 7), which demonstrates that two competing senders can form a stable system that supports full revelation. More broadly, this example can be regarded as a justification for the adversarial system.

Unlike in the simultaneous multi-sender Bayesian model (Gentzkow and Kamenica 2017b), full revelation is not always an equilibrium. As discussed in Section 9, games could have unique partial revelation in equilibrium. The relationship between a sequential game and its corresponding simultaneous game is that when the sequential equilibrium is unique,[5] no simultaneous equilibrium is less informative than the sequential one (Li and Norman 2017b).

In this paper, the informativeness is ordered by Blackwell's order. For the purpose of comparing informativeness in an intuitive way, I adopt a geometric version of Blackwell's order in regard to posterior dispersion, as will be shown by Theorem 5. Roughly speaking, it is shown that more informative signaling rules lead to more dispersed posteriors under some condition. For example, the null signal and full revealing signals lead to posteriors concentrated at the prior and vertices, respectively.

The remainder of this paper is organized as follows. Section 2 summarizes the existing literature in this area. Section 3 presents the basic model, explains the information environment, and introduces mathematical preliminaries. Section 4 explains the recursive concavification method and applies it to characterize SPE paths. Section 5 characterizes the MPE. Section 6 proves the existence of the Silent Equilibrium. Section 7 discusses the equilibrium in a zero sum game. Section 8 explores a geometric way of comparing informativeness under Blackwell's order. Section 9 discusses the relationship with the coordinated simultaneous multi-sender Bayesian persuasion model. Section 10 explores the effects of the order of persuasion. Section 11 concludes.

[5]Any equilibrium can be constructed using multiple equivalent signaling rules. So uniqueness can not be defined in regard to strategies. Instead, uniqueness in this paper means that the distribution of induced posteriors and associated receiver's optimal actions are unique.

2 Literature Review

This paper lies in the domain of Bayesian Persuasion with multiple senders. Kamenica and Gentzkow (2011) lay out a formal framework of the Bayesian persuasion game with a single sender and characterize the Sender-preferred SPE[6] by the concavification method (Aumann and Maschler 1995). Gentzkow and Kamenica (2017a) (2017b) analyze the case in which multiple senders move simultaneously and conclude that competition increases disclosure of information. Li and Norman (2017a) attract attention to the Bayesian persuasion game with multiple senders who move sequentially by constructing a counter example to Kamenica and Gentzkow (2017a) (2017b)'s general conclusion that competition improves transparency. They also propose the Markov Perfect Equilibrium as an important equilibrium concept for the sequential Bayesian persuasion game, where "belief" serves as the state variable (Li and Norman 2017a, 2017b; Ely 2017).

Li and Norman (2017b) independently address the same problem as this paper and conclude the same result that the Silent Equilibrium (the "one-step" equilibrium in their language) always exists. However, there are still many differences. First, they express coordinated signaling rules in the form of a partition framework (Li and Norman 2017a, 2017b; Kamenica and Gentzkow 2017b), instead of the traditional definition set forth in this paper. Second, they assume a finite action space so that value functions can be divided into finite polytopes; the present paper, in contrast, allows a compact action set and smooth value functions. Third, they frame the question as a linear programming problem and characterize the MPE, while this paper develops the recursive concavification method to characterize both the SPE and MPE. Fourth, this paper discusses several related topics of zero-sum games, a geometric Blackwell's order, and the order of perusasion.

Harris (1985), Harris, Reny and Robson (1995) and Hellwig, Leininger, Reny and Robson (1990) discuss the existence and characterization of SPE of infinite games, which class of games the Bayesian persuasion model belongs to. Following Harris (1985), I characterize the SPE paths by translating his method into a geometric way, i.e., the recursive concavification method. Ely (2017) uses a similar method as the recursive concavification method to solve a dynamic persuasion mechanism, where one sender controls the receiver's belief over a stochastic process. To solve a Bellman equation with a concave closure on the right side, he approximates it by repeatedly concavifying a value function. In contrast, this paper concavifies a group of value

[6] In this equilibrium, the receiver takes an action in favor of the sender when she is indifferent.

correspondences[7] of all senders.

This paper also broadly relates to the literature in communication games with multiple senders or receivers. Krishna and Morgan (2001), Battaglini (2002) and Ambrus and Takahashi (2008) extend Crawford and Sobel (1982) into multi-sender case and discuss the conditions under which full revealing equilibrium can be induced. Similar to Prop. 10, Krishna and Morgan (2001) have a result that it is beneficial for the receiver to consult both senders with interests biased in opposite directions. Milgrom and Roberts (1986) and Forges and Koessler (2008) study information disclosure models, introduced by Milgrom (1981) and Grossman (1981), with multiple senders. Bergemann and Morris (2016) introduce a new equilibrium concept, Bayesian Correlated Equilibrium, which can be interpreted as the equilibrium of a persuasion model with one sender and multiple receivers.

3 Model

3.1 Basic Setup

This section lays out a formal framework of a multi-sender sequential Bayesian persuasion game. The states of the world are $\Omega = \{\omega_1, \ldots, \omega_l\}$, $l \in \mathbb{N}$. There are T senders and 1 Receiver. Each sender has access to a set of costless signaling rules, Π. A signaling rule is a mapping from the state space to distributions over the signal space $\pi : \Omega \to \Delta(S)$, where the finite signal space is $S = \{s^1, \ldots, s^m\}$, $m \in \mathbb{N}$.[8] Equivalently, $\Pi = \prod_{\omega=1}^{l} \Delta(S)$. Specifically, let S_t denote Sender t's signal space and s_t the signal realized in period t.

Players have a common prior μ_0 in $\Delta(\Omega)$. μ_t represents the updated posterior belief after period t, for any $t = 1, \ldots, T$. The Receiver takes an action $a \in A$, where A is compact. Utility functions of all players depend on the state of the world and the action taken by the Receiver. Additionally, I impose continuity on the senders' utility functions, $v^t(a, \omega), t = 1, \ldots, T$, and the Receiver's utility function, $u(a, \omega)$.

The timing of the game is as follows:

Date 0 Nature picks a state ω according to the probability μ_0. ω is unknown to all players.

[7] As will be shown by Theorem 1, "concavifying correspondences" means imposing concave closure over the minimum of the correspondences, which becomes a criterion for selecting SPE paths.

[8] The assumption of finite signal space is made WLOG for expositional simplification. As Prop. 6 will show, the equilibrium values are invariant in the cases of finite and infinite signal spaces.

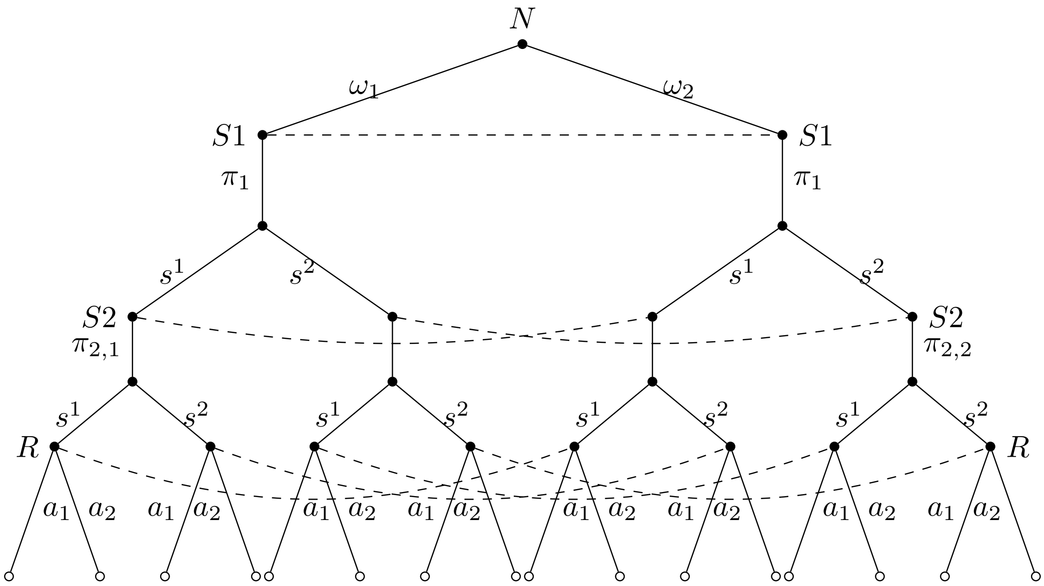


Figure 2: The game tree of a path in a Bayesian persuasion model with 2 senders

Date 1 Sender 1 chooses $\pi_1 \in \Pi$ which generates $s_1 \in S$.

Date 2 After observing $\{\pi_1, s_1\}$, Sender 2 chooses $\pi_2 \in \Pi$ which generates $s_2 \in S$.

...

Date T After observing $\{\pi_i,\, s_i\}_{i=1}^{T-1}$, Sender T chooses $\pi_T \in \Pi$ which generates $s_T \in S$.

Date $T+1$ After observing $\{\pi_i,\, s_i\}_{i=1}^{T}$, the Receiver makes a decision $a \in A$.

There are two types of information sets, one at which a sender has chosen the signaling rule but has not sent the signal, and the other at which both the signaling rule and the signal have been sent (denoted by nodes labeled "$S1$", "$S2$" and "R"). The former leads to fixed simple probability distributions over the second type of information sets without any strategic interaction. So it suffices to focus on the second type. Combining a pair of a signaling rule and its signal into a period of history, a history up until the end of date t is denoted by $e_t = (\pi_1, s_1, \ldots, \pi_t, s_t)$. E_t is the set of histories including e_t. For $t = 1, \ldots, T$, Sender t's strategy is a mapping $\sigma_t : E_{t-1} \to \Pi$ and the Receiver's strategy is $\rho : E_T \to A$. The set of Sender t's strategies is Σ_t, for $\forall t$, and the Receiver's strategy set is Σ_{T+1}.

In this paper, I focus on pure-strategy equilibrium. However, the result equivalently applies to the mixed strategy equilibrium with finite actions. Because for senders, there is no difference between using mixed strategies and pure strategies;[9] for the Receiver, the mixed strategy set with finite actions is compact.

Next, let me clarify the meaning of "subgame" in this paper. A subgame is defined to begin at an information set for each history $e_t \in E_t$, $t = 1, \ldots, T$, instead of at a single decision node as SPE usually requires. The rationale depends on that in each history and its corresponding information set, players observe, equivalently, a sequence of "experiments" and their realizations, so that it is possible to prescribe beliefs over all information sets by Bayes rule, no matter whether they are on or off the equilibrium path. With such a fixed belief system, the sequential rationality is satisfied by the players' maximizing their expected payoffs in all "subgames" originating from those information sets. Throughout this paper, I refer to SPE as with subgames in this sense.[10]

I denote the subgame starting at history e_t by $\Gamma(e_t)$, and denote the unique belief over the information set associated with e_t by $\mu_t(e_t)$. One SPE of $\Gamma(e_t)$ is represented as $\gamma(e_t)$ which consists of a strategy profile of subsequent players $\{\sigma_{t+1}, \ldots, \sigma_T, \rho\}$. One SPE of the whole game is then $\gamma(e_0)$. The set of SPE paths of $\Gamma(e_t)$ is denoted by $\bar{\Gamma}(e_t)$, and the equilibrium path of $\gamma(e_t)$ is written as $\bar{\gamma}(e_t)$. With these preliminaries at hand, I can give the formal definition of SPE:

Definition 1. *A Subgame Perfect Equilibrium of this game, $\gamma(e_0)$, is a strategy profile $\{\sigma_1, \ldots, \sigma_T, \rho\}$ such that for any $t = 0, \ldots, T$, $e_t \in E_t$ and $\sigma'_h \in \Sigma_h, h \in \{t+1, \ldots, T\}$, it satisfies:*

$$E_{\mu_t(e_t)}[v^h(\sigma_h, \sigma_{-h})] \geq E_{\mu_t(e_t)}[v^h(\sigma'_h, \sigma_{-h})]$$

And for all $e_T \in E_T$ and $a' \in A$,

$$E_{\mu_T(e_T)}[u(\rho(e_T), \omega)] \geq E_{\mu_T(e_T)}[u(a', \omega)]$$

Of special importance is a collection of values $\{V_t^k(\mu)\}_{\forall t, \forall k}$, $t \in \{0, \ldots, T\}$. $k \in \{1, \ldots, T\}$. $V_t^k : \Delta(\Omega) \Rightarrow \mathbb{R}$ is defined as $V_t^k(\mu) = \{v \in \mathbb{R} | v$ *is the payoff*

[9] Kamenica and Gentzkow 2011, footnote 3.

[10] This variation of SPE agrees with Li and Norman (2017a) (2017b). They propose SPE as a proper concept because there is no private information and players don't have to update beliefs about others' types.

for Sender k in some $\bar{\gamma}_t(e_t)$ where $\mu_t(e_t) = \mu, e_t \in E_t\}$, for $\forall t, \forall k, \forall \mu$. They represent the set of equilibrium payoffs for sender k after period t at a certain posterior belief μ. For example, if $v \in V_t^k(\mu)$, it means there is an SPE in which Sender k holds a belief μ after period t and his expected payoff is v. In Section 4, I use $\bar{V}$ and $\underline{V}$ to represent the maximum and minimum selections of a sender's value. I also use $V^k : \bar{\Gamma}(e_t) \to \mathbb{R}$, for $\forall t$, to denote Sender k's expected payoff on SPE paths.

3.2 Coordinated Information Structure

The information structure is determined by signaling rules designed by senders. The signaling rule is equivalent to the "experiment" defined in Blackwell (1951) (1953) which provides the Receiver with a more "informative" environment. If senders impose signaling rules individually, the probabilities of signals are independent with each other. However, in this model senders move sequentially and subsequent senders can vary their signaling rules over previous realizations, so each strategy profile specifies a sequence of signaling rules as $\{\pi_1, \{\pi_{2,s}\}_{s\in S_1}, \dots, \{\pi_{T,s}\}_{s\in S_1\times\dots\times S_{T-1}}\}$.

In the coordinated information structure, senders are able to correlate the probabilities of their own signals with that of previous senders'. Denote a coordinated signaling rule by π^c.

Definition 2. $\{\pi_1^c, \dots, \pi_T^c\}$ *are coordinated signaling rules if for* $\forall t$, $\forall\{\pi_1^c, \dots, \pi_{t-1}^c\}$ *and their realizations* $\{s_1, \dots, s_{t-1}\}$, π_t^c *is a mapping:* $\Omega \to \Delta(\prod_{i=1}^t S_i)$ *consistent with the probability distribution* $\Delta(\prod_{i=1}^{t-1} S_i)$ *prescribed by* $\{\pi_1^c, \dots, \pi_{t-1}^c\}$, *i.e.,* $\sum_{j=1}^m P(s_1, \dots, s_{t-1}, s_t^j|\omega) = P(s_1, \dots, s_{t-1}|\omega)$, *for* $\forall\omega$, $\forall\{s_1, \dots, s_{t-1}\} \in S_1 \times \dots \times S_{t-1}$.

With a coordinated signaling rule a sender can unilaterally deviate to any information structure weakly more informative than the existing one (Gentzkow and Kamenica 2017b), which plays an important role in the proof of Theorem 4 and Prop. 12.

Prop. 1 reveals the equivalence relation between information structures generated by $\{\pi_1, \{\pi_{2,s}\}_{s\in S_1}, \dots, \{\pi_{T,s}\}_{s\in S_{T-1}}\}$ and $\{\pi_1^c, \dots, \pi_T^c\}$. It also clarifies the coordinated information structure embedded in this game and estabilishes connections with other work in multi-sender persuasion games (Li and Norman 2017a, 2017b; Gentzkow and Kamenica 2017b). It shows that the analysis in this paper also applies to an alternative game in which senders move sequentially and design coordinated signaling rules.

Proposition 1. *The information structure given by independent signaling rules* $\{\pi_1, \{\pi_{2,s}\}_{s\in S_1}, \dots, \{\pi_{T,s}\}_{s\in S_1\times\dots\times S_{T-1}}\}$ *is equal to that generated by some* *coordinated signaling rules* $\{\pi_1^c, \dots, \pi_T^c\}$*. The converse is true.*

Proof. See appendix. □

3.3 Preliminaries

This part mentions a series of knowledge about Bayesian plausibility, some facts in convex analysis, and Blackwell's order, that form the mathematical basis of following sections. In addition, I introduce notation for expositional simplification.

3.3.1 Signaling rules and posteriors

Each signal s^j from a signaling rule π leads to a posterior ϕ_j by Bayes Law. These posteriors, $\{\phi_j\}_{j=1}^m$, and the probabilities of their associated signals, $\{p(\phi_j)\}_{j=1}^m$, construct a distribution of distributions that summarizes all information content of a signaling rule. The relationship between signaling rules and their posteriors is called *Bayesian plausibility*.[11]

Definition 3. *(Bayesian plausibility) For any* $\pi \in \Pi$ *and its posteriors* $\{\phi_j\}_{j=1}^m$, $\mu = \sum_{j=1}^m p(\phi_j)\phi_j$, $p(\phi_j) \in [0,1]$, *for* $\forall j$*. Conversely, any Bayesian plausible distributed posteriors are outcomes of some signaling rule.*

This equivalence relation allows me to define a signaling rule indirectly as its posteriors. Specifically, I refer $\langle \phi_1, \dots, \phi_m \rangle_\mu$ to a signaling rule that generates a distribution of posteriors $\{\phi_1, \dots, \phi_m\}$ whose expectation is μ, such that $\mu = \sum_{j=1}^m p(\phi_j)\phi_j$, $p(\phi_j) \in [0,1]$, for $\forall j$. $\{\langle \phi_1, \dots, \phi_m \rangle_\mu, \bar{\gamma}_1, \dots, \bar{\gamma}_m\}$ represents an SPE path on which a sender designs a signaling rule $\langle \phi_1, \dots, \phi_m \rangle_\mu$, and his followers stay on an SPE path, $\bar{\gamma}_j$, receiving his signal s^j, for $\forall j$. Furthermore, I denote by $\langle \phi_1^n, \dots, \phi_m^n \rangle_\mu \to \langle \phi_1, \dots, \phi_m \rangle_\mu$ the convergence of a sequence of signaling rules, where posteriors $\phi_j^n \to \phi_j$ and coefficients $p(\phi_j^n) \to p(\phi_j)$, for $\forall j$. I should point out that Bayesian plausibility is preserved in the limit, i.e., $\langle \phi_1, \dots, \phi_m \rangle_\mu$ still represents a signaling rule. Notice that $p(\phi_j)$ could be zero that makes the corresponding signal redundant, but it is a convenient expression in Section 4.

3.3.2 Convex Analysis

Most proofs draw on useful results from convex analysis, so it is necessary to point them out at this stage. The first is a key concept that has been widely used in economics to identify the potential rents of persuasion for senders,

[11] Referring to Kamenica and Gentzkow (2011), Prop. 1.

that is, the concave closure. Mathematically, there are two equivalent definitions.[12] The most well known definition describes it as the lowest concave function that is not lower than a specific function. Another definition, however, is given by Eq. (1). Though complicated, it perfectly fits the context of a persuasion game, and is adopted as the main definition in this paper.

$$cl(f(x)) = \sup \Big\{ \sum_{j=1}^{m} \alpha_j f(x_j) | \boldsymbol{\alpha} \in \Delta_m, \sum_{j=1}^{m} \alpha_j x_j = x, m = 1, 2, \dots \Big\} \quad (1)$$

Each component of the equation contains a corresponding economic meaning. The objective function f refers to a value function of a sender, the set $\{x_j\}_{j=1}^m$ a distribution of posteriors of a singaling rule, $\{\alpha_j\}_1^m$ the probabilities of these signals. In addition, the restriction "$\boldsymbol{\alpha} \in \Delta_m, \sum_{j=1}^m \alpha_j x_j = x$" implies the satisfication of *Bayesian plausibility.* This definition is central to the recursive concavification method in Sections 4 and 5.

On a belief space $\Delta(\Omega)$, one posterior belief is a point and the convex hull of a set of posteriors $\{\phi_1, \dots, \phi_m\}$ is denoted by $conv\{\phi_1, \dots, \phi_m\}$. A regularity in regard to convex hulls, which determines uniqueness in Section 8, is Carathéodory's Theorem.[13]

Carathéodory's Theorem

For any set $S \subset \mathbb{R}^n$ and $x \in convS$, x can be represented as a convex combination of $n+1$ elements of S.

3.3.3 Blackwell's Order

Adopted as the measure of informativeness, Blackwell's order is a *partial order* on the space of signaling rules. A more informative signaling rule creates a better environment for a decision maker to achieve higher expected payoffs. Specifically, the meaning of Blackwell's order is given by Definition 4 and Definition 5, both of which are equivalent based on Blackwell's theorem.

Suppose there are two signaling rules $\pi = \langle \lambda_1, \dots, \lambda_{m_1} \rangle_{\mu_0}$ and $\pi' = \langle \mu_1, \dots, \mu_{m_2} \rangle_{\mu_0}$, i.e. $\mu_0 = p_1\lambda_1 + \dots + p_{m_1}\lambda_{m_1}$ and $\mu_0 = q_1\mu_1 + \dots + q_{m_2}\mu_{m_2}$. The signal generating mechanisms can be described as two matrices $A_{m_1 \times l}$ and $B_{m_2 \times l}$, where the (i, j)th entry represents the probability of generating the i-th signal at state j, i.e. $a_{ij} = P(s^i|\omega_j)$ and $b_{ij} = P(s^{i'}|\omega_j)$ where s^i induces λ_i and $s^{i'}$ induces μ_i. Given the Receiver's utility function u, her value for a posterior μ is defined as $V_u : \Delta(\Omega) \to \mathbb{R}$, $V_u(\mu) =$

[12] Hiriart-Urruty, Jean-Baptiste, and Claude Lemaréchal, "Fundamentals of Convex Analysis," pp.99, Proposition 2.5.1.

[13] Ewald, Günter, "Combinatorial convexity and algebraic geometry," Section 2.3, pp.10.

$\max_{a\in A} E_\mu[u(a,\omega)]$. Her value for a signaling rule $\pi = \langle \lambda_1, \ldots, \lambda_{m_1}\rangle_{\mu_0}$ is $\bar{V}_u : \Pi \to \mathbb{R}$, $\bar{V}_u(\pi) = p_1 V_u(\lambda_1) + \cdots + p_{m_1} V_u(\lambda_{m_1})$. An *action rule* is a mapping from the signal space to the action space, $f : S \to A$, and the set of action rules is $\mathcal{F}$. Defining the expected payoff of an action rule f at state i as $v_i(f,\pi) = \sum_{j=1}^{m} \pi(s_j|\omega_i) u(f(s_j), \omega_i)$, $\forall i = 1, \ldots, l$, the range of the payoff vector is denoted by $D_\pi = \{(v_1(f,\pi), \ldots, v_l(f,\pi)) | f \in \mathcal{F}\}$.

Before introducing Blackwell's theorem, I define *sufficiency* and *informativeness* as two relations among signaling rules.

Definition 4. *(Sufficiency) If π' is sufficient for π, written $\pi' \succ \pi$, then there exists a matrix $C_{m_1\times m_2}, c_{ij} \geq 0, \sum_i c_{ij} = 1, \forall i,j$, such that $A = CB$.*

Definition 5. *(Informativeness) If π' is more informative than π, written $\pi' \supset \pi$, then $D_{\pi'} \supset D_\pi$, for any utility function u and any compact action set A.*

Blackwell (1951)(1953) demonstrates the equivalence between the statistical relation (sufficiency) and the Bayesian decision relation (informativeness), regardless of the common prior, the action set and the payoff structure. This model, however, is endowed with a common prior that facilitates a geometric expression of Blackwell's order, as discussed in Section 8.

Blackwell's Theorem $\pi' \succ \pi$ *iff* $\pi' \supset \pi$.

4 Subgame Perfect Equilibrium

This section provides the characterization of the set of SPE paths. In this extensive game, senders' persuasion forms a process of strategic information revelation, where the realized beliefs become the "priors" of following subgames. Not only does the Receiver, but also the senders depend their actions on those realized beliefs, which gives rise to a continuum of values for players conditional on each posterior belief of each period, i.e. $\{V_t^k\}_{\forall t,\forall k}$.

To resolve multiplicity, I adopt Harris (1985)'s method to characterize the SPE paths recursively. This method relies on imposing the lower bound on the SPE payoff to the active sender in each period, called Harris condition in this paper, to identify the sequence of signaling rules that constitute SPE paths. Furthermore, Harris condition is sufficient for selecting SPE and takes a geometric form of the concave closure of the minimum value function.

In application, the set of SPE paths is characterized backward. It is immediate to figure out the Receiver's behavior given information from all sources, which leads to the set of all senders' equilibrium payoffs dependent on the final beliefs, i.e. $\{V_T^k\}_{k=1}^T$. By applying Harris condition on Sender T, one can solve his equilibrium signaling σ_T^S and its payoff implications to other senders conditional on the belief of the second to the last period, i.e., $\{V_{T-1}^k\}_{k=1}^T$. Repeating this process yields senders' strategies and equilibrium payoffs at each stage until the set of SPE paths is achieved.

4.1 Harris (1985)

Implicit in any SPE is credible punishment for all potential deviations, which is reflected in Harris condition. This condition measures the highest payoff a sender can defend himself against the most credible "punishment" from his followers. It is necessary and sufficient that senders receive payoffs no less than Harris condition in equilibrium.

Harris condition can be found in two steps. First, for any $t \in \{1, \dots, T\}$, suppose $\bar{\Gamma}_t(e_t)$ are known for $\forall e_t \in E_t$, calculate the worst equilibrium payoff for Sender t when he takes a signaling rule $\pi_t := \langle \phi_1, \dots, \phi_m \rangle_{\mu_{t-1}(e_{t-1})}$, denoted by $b(\pi_t)$.

$$
\begin{aligned}
b(\pi_t) =& \min \Big\{ \sum_{j=1}^{m} p(\phi_j) v_j \mid v_j \in V_t^t(\mu_t(e_{t-1}, \pi_t, s^j)), \forall j \Big\} \\
=& \sum_{j=1}^{m} p(\phi_j) \min\{V_t^t(\mu_t(e_{t-1}, \pi_t, s^j))\} \\
=& \sum_{j=1}^{m} p(\phi_j) \underline{V}_t^t(\mu_t(e_{t-1}, \pi_t, s^j))
\end{aligned} \tag{2}
$$

This is well defined because all value correspondences have closed graphs by Prop. 2 and Prop. 5, and thus the minimum can be achieved. The second step is to find out the supremum of these worst payoffs, $\sup\{b(\pi')|\pi' \in \Pi\}$ (Harris condition). [14]

Formally, suppose $\bar{\gamma}_j$ is an SPE path in $\bar{\Gamma}((e_{t-1}, \pi_t, s^j))$ for $\forall j$, then $\{\pi_t, \{\bar{\gamma}_1, \dots, \bar{\gamma}_m\}\}$ is one of the SPE paths in $\bar{\Gamma}_{t-1}(e_{t-1})$ if and only if Sender t's expected payoff from it is no less than $\sup\{b(\pi')|\pi' \in \Pi\}$. Because of the definition of the concave closure in Eq. (1) and the equivalence between signaling rules and Bayesian plausible distributed posteriors, Harris condition takes another geometric form: the concave closure of the minimal values.

[14] Referring to Harris (1985), Sec. 4.1.

$$
\begin{aligned}
&\sup\{b(\pi')|\pi' \in \Pi\} \\
&= \sup\Big\{\sum_{j=1}^{m} p(\phi'_j)\underline{V}_t^t(\mu_t(e_{t-1}, \pi_t, s^j)) \mid \forall\langle\phi'_1, \ldots, \phi'_m\rangle_{\mu_{t-1}(e_{t-1})}\Big\} \\
&= cl(\underline{V}_t^t)(\mu_{t-1}(e_{t-1}))
\end{aligned} \tag{3}
$$

Eq. (3) establishes a connection between the condition and the concavification method, with which one can pin down the sequence of values $\{V_t^k\}_{\forall t, \forall k}$ as well as the set of SPE paths recursively:

$$
\{V_T^t\}_{t=1}^T \Rightarrow cl(\underline{V}_T^T) \Rightarrow \{V_{T-1}^t\}_{t=1}^T \Rightarrow cl(\underline{V}_{T-1}^{T-1}) \Rightarrow \ldots \Rightarrow \{V_0^t\}_{t=1}^T
$$

4.2 Recursion

4.2.1 Period $T + 1$

At the beginning of period $T + 1$, the Receiver has observed all signaling rules and signals from senders and formed a new belief μ_T, based on which she optimizes her expected utility by choosing an action. The optimal action set is $A^*(\mu_T) = \{a \in A | a \in \arg\max_{a \in A} E_{\mu_T}(u(a, \omega))\}$. Because of continuity of u and compactness of the domain, by the Maximum Theorem, A^* is non-empty valued and has a closed graph. The next proposition establishes a more significant result that V_T^t is non-empty valued and has a closed graph for $t = 1, \ldots, T$.

Proposition 2. *For each* $t = 1, \ldots, T$, $V_T^t(\mu_T) = \{E_{\mu_T}(v^t(a, \omega)) | a \in A^*(\mu_T)\}$. V_T^t *is non-empty valued and has a closed graph.*

Proof. For $\forall\{\mu^n\} \to \mu$ and $\{v^n | v^n \in V_T^t(\mu^n)\} \to v$, $\exists\{a^n\} \subset A$ s.t. $a^n \in A^*(\mu^n)$, $E_{\mu^n}(v^t(a^n, \omega)) = v^n$.

Because A is compact and A^* has a closed graph, $\exists\{a^{n_k}\} \subset \{a^n\} \to a^* \in A^*(\mu)$. Therefore $E_\mu(v^t(a^*, \omega)) \in V_T^t(\mu)$.

Because $\{a^{n_k}\} \to a^*$ and v^t continuous,

$$
v = \lim_{n\to\infty} v^n = \lim_{k\to\infty} E_{\mu^{n_k}}[v^t(a^{n_k}, \omega)] = E_\mu[v^t(a^*, \omega)] \in V_T^t(\mu)
$$

□

Prop. 2 shows that the condition for the application of Harris condition is satisfied. The set of SPE paths, $\bar{\Gamma}_T(e_T)$, for $\forall e_T \in E_T$, is

$$
\bar{\Gamma}_T(e_T) = A^*(\mu_T(e_T)) \tag{4}
$$

4.2.2 Period T

Subgames of period T consisting of the Receiver and Sender T resemble single-sender persuasion games. Equivalently, one can solve them by the concavification method with Harris condition.

Proposition 3. *For* $\forall \mu \in \Delta(\Omega)$, $v \in V^T_{T-1}(\mu)$ *if and only if* $\exists \langle \theta_1, \dots, \theta_m \rangle_\mu$ *s.t.* $v = \sum_{j=1}^m p(\theta_j) v_j, v_j \in V^T_T(\theta_j)$ *and* $v \geqslant cl(\underline{V}^T_T)(\mu)$.

Proof. (Sufficient) Let Sender T send a null signal and the Receiver take an action that support v in a SPE. For other signaling rules Sender T designs, $\langle \phi_1, \dots, \phi_m \rangle_\mu$, let the Receiver take an optimal action that yields $\underline{V}^T_T(\phi_j)$ for Sender T upon receiving each ϕ_j, $\forall j = 1, \dots, m$. Then by the definition of $cl(\underline{V}^T_T)(\mu)$, the value for Sender T of playing any other signaling rule is dominated by $cl(\underline{V}^T_T)(\mu)$, and by v as well.

(Necessary) If $v < cl(\underline{V}^T_T)(\mu)$, by the definition of $cl(\underline{V}^T_T)$, $\exists \langle \phi_1, \dots, \phi_m \rangle_\mu$ s.t.
$\sum_{j=1}^m p(\phi_j) \underline{V}^T_T(\phi_j) > v$, i.e. the signaling rule $\langle \phi_1, \dots, \phi_m \rangle_\mu$ dominates the value v in any SPE following this strategy. □

As Harris condition is necessary and sufficient, the SPE paths are simply action profiles satisfying it.

$$\begin{aligned}\bar{\Gamma}_{T-1}(e_{T-1}) = \Big\{ & \{\pi_T, a_1, \dots, a_m\} \mid \pi_T := \langle \phi_1, \dots, \phi_m \rangle_{\mu_{T-1}(e_{T-1})}, \\ & a_j \in \bar{\Gamma}_T(e_{T-1}, \pi_T, s^j), \forall j, \text{ and} \\ & \sum_{j=1}^m p(\phi_j) V^T(\{a_j\}) \geq cl(\underline{V}^T_T)(\mu_{T-1}(e_{T-1})) \Big\} \end{aligned} \tag{5}$$

Eq. (5) identifies $\bar{\Gamma}_{T-1}(e_{T-1}), \forall e_{T-1} \in E_{T-1}$, which naturally yield the values of senders in period $T-1$. As mentioned before, all subgames at this period, starting with the same "prior," share the same set of SPE paths, which allows me to group histories by beliefs and define value correspondences on the belief space. Therefore, $\{V^t_{T-1}\}^T_{t=1}$ incorporate all payoffs supported in SPE, i.e., $\bar{\Gamma}_{T-1}(e_{T-1}), \forall e_{T-1} \in E_{T-1}$.

$$\begin{aligned} V^t_{T-1}(\mu) = \Big\{ & V^t(\bar{\gamma}) \mid \exists e_{T-1} \in E_{T-1}, \mu_{T-1}(e_{T-1}) = \mu, \\ & \bar{\gamma} \in \bar{\Gamma}_{T-1}(e_{T-1}) \Big\} \end{aligned} \tag{6}$$

The existence of SPE is also guaranteed. In contrast with the minimum of the value for characterization, I use the maximum of the value to show the existence. Due to the closedness of V_T^T, the maximum of the value correspondence is USC. With this property, those signaling rules that support the concave closure of the maximum of the value are equilibrium strategies by Prop. 3, as they also generate payoffs no less than Harris condition.

Proposition 4. *For $\forall \mu \in \Delta(\Omega)$, and any $v = cl(\bar{V}_T^T)(\mu)$, v is an equilibrium payoff for Sender T in a SPE $\gamma(e_{T-1})$ s.t. $\mu_{T-1}(e_{T-1}) = \mu$.*

Proof. By the definition of $cl(\bar{V}_T^T)$, $\exists\{\langle \theta_1^n, \ldots, \theta_m^n \rangle_{\mu_{T-1}}\}$, such that $\sum_{k=1}^m p_k^n \bar{V}_T^T(\theta_k^n) \to v$. Because $\Delta(\Omega)$ is compact, $\exists \langle \theta_1^{n_k}, \ldots, \theta_m^{n_k} \rangle_{\mu_{T-1}} \to \langle \theta_1^*, \ldots, \theta_m^* \rangle_{\mu_{T-1}}$.

Because V_T^T has a closed graph, $\bar{V}_T^T$ is USC, i.e. $\bar{V}_T^T(\theta_j^*) \geqslant \lim \bar{V}_T^T(\theta_j^{n_k})$, for $\forall j$. Therefore, $\sum_{j=1}^m p_j^* \bar{V}_T^T(\theta_j^*) \geqslant \lim \sum_{j=1}^m p_j^{n_k} \bar{V}_T^T(\theta_j^{n_k}) = v = cl(\bar{V}_T^T)(\mu_{T-1}) \geqslant cl(\underline{V}_T^T)(\mu_{T-1})$. Also, as V_T^T has closed graph, $\bar{V}_T^T(\theta_j^*) \in V_T^T(\theta_j^*)$, so $\langle \theta_1^*, \ldots, \theta_m^* \rangle_{\mu_{T-1}}$ is an equilibrium play of Sender T at period T. □

4.2.3 Period t, $t < T$

This part extends the analysis to earlier stages by repeating the above argument. Similar results to Prop. 3 and 4 hold for $\{V_t^t\}_{t=1}^T$, for $\forall t$. Take period $T-1$ for instance. With the knowledge of V_{T-1}^{T-1}, Sender $T-1$ can evaluate the consequence of a certain signaling rule as well as Sender T does. Thus, the set of SPE paths, $\bar{\Gamma}_{T-2}$, is derived from the same condition except for using a new value correspondence, V_{T-1}^{T-1}. Then one can obtain Sender $T-2$'s value V_{T-2}^{T-2} he cares about and continue to solve $\bar{\Gamma}_{T-3}$, so on and so forth.

Theorem 1. *For $\forall t$, $\forall \mu \in \Delta(\Omega)$, $v \in V_t^{t+1}(\mu)$ if and only if $\exists \langle \theta_1, \ldots, \theta_m \rangle_\mu$ s.t. $v = \sum_{j=1}^m p(\theta_j) v_j, v_j \in V_{t+1}^{t+1}(\theta_j)$ and $v \geqslant cl(\underline{V}_{t+1}^{t+1})(\mu)$.*

Proof. Similar to the proof of Prop. 3. □

Theorem 2. *For $\forall t$, $\forall \mu \in \Delta(\Omega)$, and any $v = cl(\bar{V}_{t+1}^{t+1})(\mu)$, v is an equilibrium payoff for Sender $t+1$ in a SPE $\gamma_t(e_t)$ s.t. $\mu_t(e_t) = \mu$.*

Proof. Similar to the proof of Prop. 4. □

Theorems 3 and 4 suggest that the set of SPE paths, $\bar{\Gamma}_{t-1}(e_{t-1})$, and its corresponding set of equilibrium payoffs, V_{t-1}^k, for $\forall k$, can be obtained by formulas similar with Eq. (5)(6). For $\forall e_{t-1} \in E_{t-1}$,

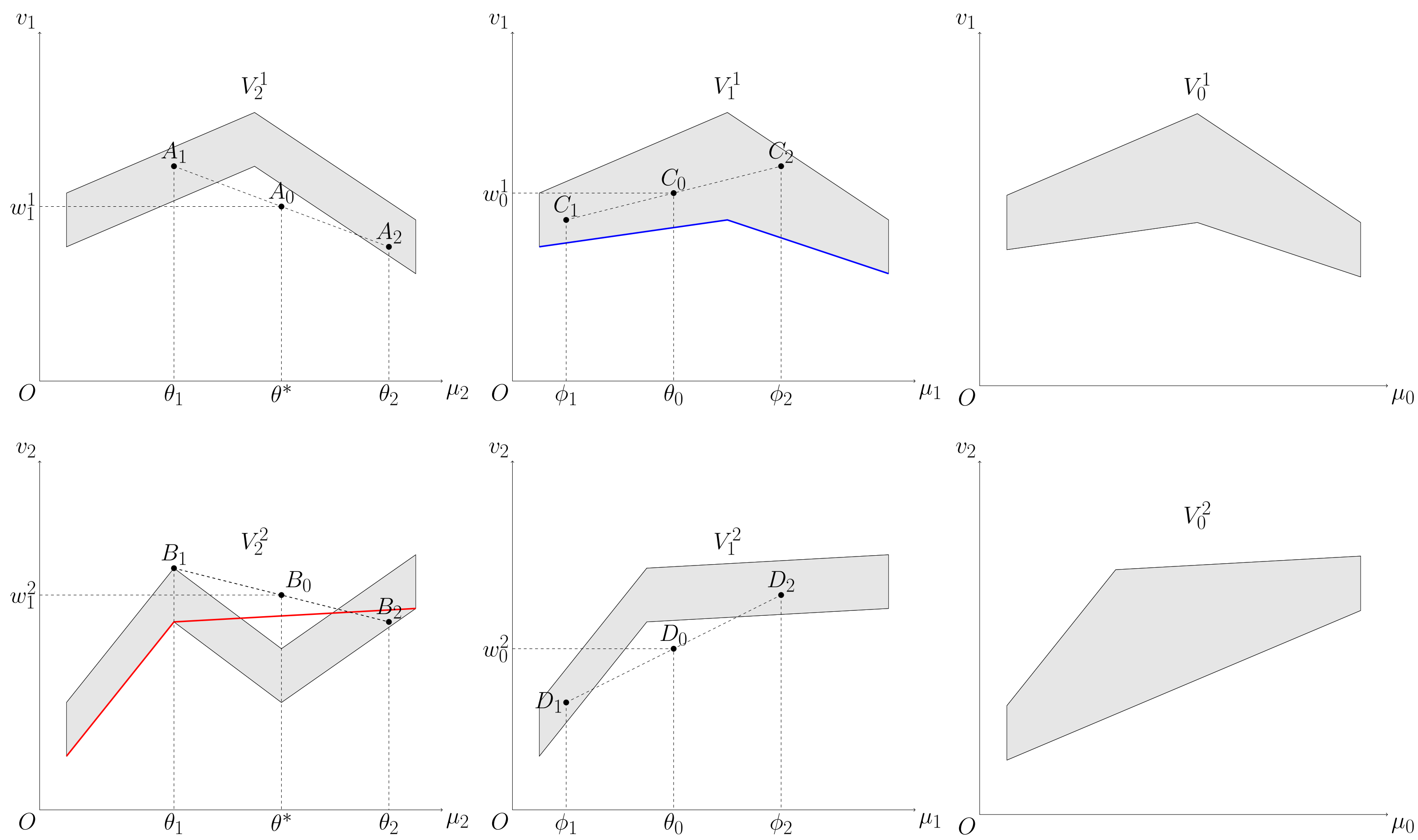


Figure 3: A Geometric Illustration of an SPE

$$\bar{\Gamma}_{t-1}(e_{t-1}) = \Big\{\{\pi_t, \gamma_1, \ldots, \gamma_m\} \mid \pi_t := \langle \phi_1, \ldots, \phi_m \rangle_{\mu_{k-1}(e_{k-1})}, \\ \bar{\gamma}_j \in \bar{\Gamma}_t(e_{t-1}, \pi_t, s^j), \forall j, \text{ and} \\ \sum_{j=1}^{m} p(\phi_j) V^t(\{\bar{\gamma}_j\}) \geq cl(\underset{t}{V^t})(\mu_{t-1}(e_{t-1}))\Big\} \tag{7}$$

$$V^k_{t-1}(\mu) = \Big\{V^k(\bar{\gamma}) \mid \exists e_{t-1} \in E_{t-1}, \mu_{t-1}(e_{t-1}) = \mu, \\ \bar{\gamma} \in \bar{\Gamma}_{t-1}(e_{t-1})\Big\} \tag{8}$$

Now I have given the set of SPE paths and defined a sequence of values for each sender at each period, the remaining challenge is to show that $\{V^k_t\}_{\forall t, \forall k}$ are closed. This desirable mathematical property builds on many features of this game, including the continuous value functions, the compact action space, and the compact set of signaling rules, Π. In addition, finiteness is another important condition approached by focusing on equilibrium paths including only finite action profiles. In summary, all these properties such as continuity, compactness and finiteness, lead to the closedness of the value correspondences.

For any V^k_t, any converging sequence of value points has a limiting point supported by the limiting SPE path. This Upper hemi-continuity property is similar with the result in Hellwig, Leininger, Reny and Robson (1990), where they use SPE paths of "nearby" finite games to approximate the SPE path of a targeting infinite game.

Proposition 5. *V^k_t have closed graphs, for $\forall t, \forall k$.*

Proof. See Appendix. □

Finally, the assumption of a finite signal space can be relaxed. Denote a persuasion game with an infinite signal space by Γ_∞, and another with a sufficiently large finite signal space by Γ_0, such that $|S| \geq T + l$. The equivalence between Γ_∞ and Γ_0 is presented below.

Proposition 6. *All equilibrium values in Γ_∞ are also equilibrium values in Γ_0.*

Proof. See Appendix. □

4.3 A Geometric Illustration

Figure 3 shows how I use Harris condition to characterize the set of SPE paths in a recursive way. In the left column are both senders' equilibrium payoffs V_2^1 and V_2^2, that are defined on the final belief space, taking into account the Receiver's optimal response.

In the second period, Harris condition on Sender 2 is illustrated by the red curve. A typical signaling rule designed by Sender 2 in equilibrium, for a prior of the second period, θ^*, is to randomize between posteriors θ_1 and θ_2. Followed by the Receiver's optimal responses that generates payoffs A_1, B_1 or A_2, B_2 to both senders, this signaling rule $\langle\theta_1, \theta_2\rangle_{\theta^*}$ is justified in equilibrium if and only if Sender 2's payoff from using it is no lower than Harris condition. Satisfying this condition, Sender 2's expected payoff is contained in Sender 2's value of the second period at θ^*, i.e. $w_1^2 \in V_1^2(\theta^*)$. In parallel, Sender 1's expected payoff on this path belongs to his value of the second period, i.e., $w_1^1 \in V_1^1(\theta^*)$.

In the middle column, V_1^1 and V_2^1 consist of all equilibrium payoffs to Sender 1 and Sender 2, in subgames starting from the second period. Similarly, Sender 1's behavior can be analyzed by imposing Harris condition as the concave closure of the minimal value of V_1^1, indicated by the blue curve. For the prior of the game, θ_0, Sender 1's signaling rule $\langle\phi_1, \phi_2\rangle_{\theta_0}$, followed by the SPE paths of subsequent periods, is supported in equilibrium because Sender 1 receives expected payoff no lower than Harris condition. Sender 1 and Sender 2's payoffs on this path are therefore contained in $V_0^1(\theta_0)$ and $V_0^2(\theta_0)$, respectively. In the end, $V_0^1(\theta_0)$ and $V_0^2(\theta_0)$ summarize the equilibrium payoffs to both senders.

4.4 Ranges of Values

Figure 4 illustrates the limits of persuasion in three scenarios: a single-sender case, a multi-sender case and a case in which the player can not persuade. Suppose the sender, given the Receiver's optimal actions, has a value correspondence identified by the union of the two black curves.[15] In a single-sender case, SPE payoffs are no less than $cl(\underline{V})$ according to Theorem 1 and no larger than $cl(\bar{V})$, lying in the area A. When the sender competes with others, his payoffs may fall below $cl(\underline{V})$, but not lower than the hyperplane going across the two lowest values at vertices, denoted by h; this is his guaranteed payoff by revealing the truth. The range of payoffs in the second case is $A \cup B$. If the sender is not allowed to persuade, he may get a more detrimental result

[15]By construction, the receiver has 2 optimal actions at each μ_T.

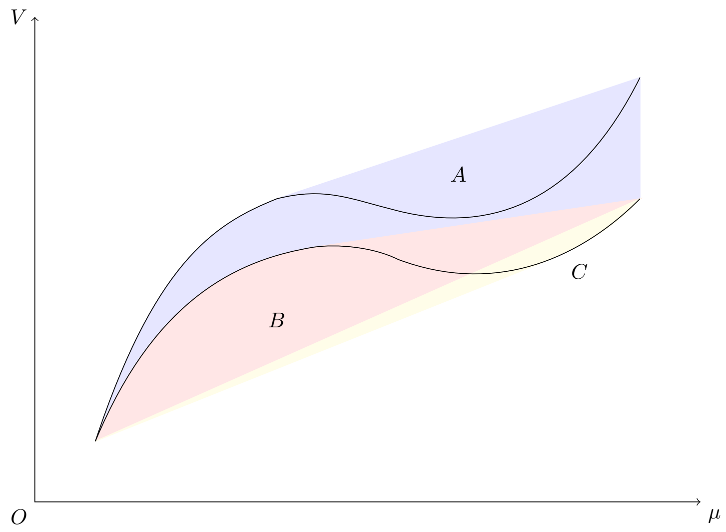


Figure 4: Ranges of values in 3 cases

in C. The lower bound of C is the convex closure of $\underline{V}$ from bottom, denoted by b. The possible equilibrium payoffs are confined in $A \cup B \cup C$, the convex hull of the value correspondence.

Proposition 7. *For any Bayesian persuasion model with one Receiver and one sender: 1) the range of SPE payoffs in a single sender game is between* $cl(\bar{V})$ *and* $cl(\underline{V})$*; 2) the range of* potential *SPE payoffs in a multi-sender sequential game is between* $cl(\bar{V})$ *and hyperplane* h*; the range of* potential *SPE payoffs when he cannot persuade is between* $cl(\bar{V})$ *and surface* b*.*

The inclusion relation between these sets sheds light on the impact of persuasion on senders' welfare. When a sender monopolizes information disclosure, he is able to receive the highest range of payoffs. When he competes with other senders, his payoff may drop down, but is still guaranteed a basic rent from revealing the truth. If he cannot persuade, his equilibrium payoff could be even lower.

4.5 Multiplicity

There could be multiple equilibria, resulting from the abundant strategy sets of both senders and the Receiver. Once the Receiver has multiple opitmal actions, those value correspondences become "thick," and Harris condition is slack in a sense that uncountably many action profiles satisfy it. Thus, the explicit form of the set of SPE could be intractable, even in a simple example

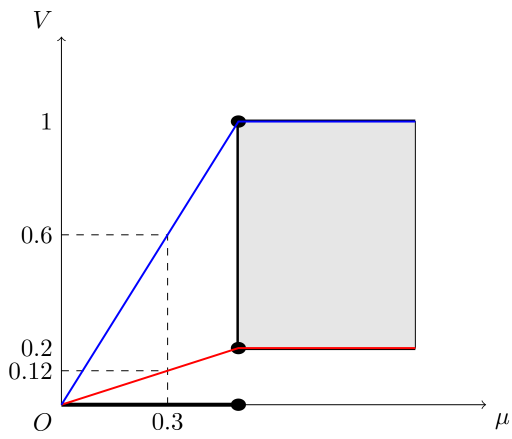


Figure 5: The judge charges a penalty

as below.

Example 2: The Judge Charges a Penalty

Suppose the judge should not only decide to "acquit" or "convict", but also impose a penalty on the defendant once he gets convicted. The penalty could be any number between 0.2 and 1 (thousand dollars), constructing a compact action set $\{0\} \cup [0.2, 1]$ ("0" refers to "acquit"). The judge's target is still raising the judgement accuracy, and her optimal strategies can be described as: $a = 0$ if $\mu < 0.5$; $a \in \{0\} \cup [0, 1]$ if $\mu = 0.5$ and $a \in [0.2, 1]$ if $\mu > 0.5$. Though the judge is herself indifferent to the penalty, the prosecutor would like as high a fine to be charged as possible. Specifically, his payoff function is $v(a, \omega) = a$ and his value is represented as a correspondence in Figure 5.

Harris condition is represented as the red curve implying that the prosecutor's expected payoff in equilibrium should be at least 0.12. The SPE is non-unique. For example, $\{\langle 0, 0.5\rangle_{0.3}, 0, a\}, \forall a \in [0.2, 1]$ is a family of SPE paths that identifies the range of SPE payoffs for the prosecutor as $[0.12, 0.6]$. However, there are still uncountably many other SPE paths, which are impossible to explicit individually.

5 The Markov Perfect Equilibrium

As shown above, the SPE might be intractable due to multiplicity, while the Markov Perfect Equilibrium can be a desirable refinement. Because in a persuasion game the only payoff relevant factor is "belief," it is the only candidate for the state variable (Li and Norman 2017a, 2017b; Ely, 2017). In MPE, players respond to realized beliefs instead of histories, in which context players only consider making full use of information regardless of how the information is revealed. This section discusses MPE from 3 respects. First of all, I characterize the MPE with a simpler version of the recursive concavification method. Next, I present a general example of the MPE. Finally, I show that the MPE payoff set is properly included by the SPE payoff set.

5.1 Characterization

Formally speaking, Sender t's Markovian stratege is defined as a mapping $\sigma_t^M : \Delta(\Omega) \to \Pi$, for $\forall t$, and the Receiver's Markovian strategy as $\rho^M : \Delta(\Omega) \to A$. The MPE shared by the class of subgames $\{\Gamma(e_t), \mu_t(e_t) = \mu\}$ is denoted by $\Gamma_t^M(\mu)$.

In MPE, the Receiver takes an action for each final belief, reducing senders' values to functions. Therefore, $cl(\bar{V})$ coincides with $cl(\underline{V})$ and Harris condition is satisfied when the expected payoff reaches the concave closure. However, existence is preserved only for value functions satisfying *achievability*, a condition that substitutes "maximum" in Eq. (1) for "supremum."

Proposition 8. *For each sender, given his value function of posterior beliefs, $f(x)$, the optimal signaling rule exists for each prior if and only if: $cl(f)(x) = \max\left\{ \sum_{j=1}^m \alpha_j f(x_j) | \alpha \in \Delta_l, \sum_{j=1}^m \alpha_j x_j = x \right\}$. When this condition is satisfied, the value function is called "achievable".*

Proof. When $cl(f)(x) = \max\left\{ \sum_{j=1}^m \alpha_j f(x_j) | \alpha \in \Delta_l, \sum_{j=1}^m \alpha_j x_j = x \right\}$, for any prior μ, there exists $\{\phi_1, \dots, \phi_m\}$ such that $\sum_{j=1}^m p(\phi_j)\phi_j = \mu$ and $\sum_{j=1}^m p(\phi_j) f(\phi_j)$
$= cl(f)(\mu)$. Obviously, $\langle \phi_1, \dots, \phi_m, \rangle_\mu$ is an optimal signaling rule.

If for some $\mu \in \Delta(\Omega)$, $\max\left\{ \sum_{j=1}^m \alpha_j f(x_j) | \alpha \in \Delta_l, \sum_{j=1}^m \alpha_j x_j = \mu \right\}$ is not well defined, which means that there is a sequence of signaling rules $\langle \phi_1^n, \dots, \phi_m^n \rangle_\mu$ whose values approximate $cl(f)(\mu)$ but do not reach the limit, which rejects the existence of an optimal signaling rule. □

The recursive concavification method resembles that for SPE. In period $T + 1$, the Receiver takes an action $\rho^M(\mu) \in A^*(\mu)$, for $\forall \mu$, that defines $\Gamma_T^M(\mu)$. For any μ, the MPE $\Gamma_T^M(\mu)$ exists and determines senders' values of the last period, $\{V_T^t(\mu)\}_{t=1}^T$.

$$V_T^t(\mu) = E_\mu[v^t(\rho^M(\mu), \omega)] \tag{9}$$

$\{V_T^t\}_{t=1}^T$ could be any selections of the value correspondences $\{V_T^t\}_{\forall t}$ without restriction, but the MPE of period $T-1$, $\Gamma_{T-1}^M(\mu)$, exists for any μ if and only if V_T^T is *achievable* according to Prop. 8. Based on the concavification method, $\sigma_T^M := \langle \phi_1(\mu), \dots, \phi_m(\mu) \rangle_\mu$ is an optimal strategy if and only if for $\forall \mu$:

$$cl(V_T^T)(\mu) = \sum_{j=1}^{m} p(\phi_j(\mu)) V_T^T(\phi_j(\mu)) \tag{10}$$

The satisfication of Eq.(10) directly implies that $\{\sigma_T^M(\mu), \rho\} \in \Gamma_{T-1}^M(\mu)$ by Prop. 8. Furthermore, this definition disentangles the contingent strategies of Sender T for all possible priors μ_{T-1}, and exhibits the influence of his persuasion on other passive senders. In parallel, $\{V_{T-1}^t\}_{\forall t}$ can be transformed from $\{V_T^t\}_{\forall t}$ by a formula,

$$V_{T-1}^t(\mu) = \left\{ \sum_{j=1}^{m} p(\phi_j(\mu)) V_T^t(\phi_j(\mu)) \mid \sigma_T^M(\mu) = \langle \phi_1(\mu), \dots, \phi_m(\mu) \rangle_\mu \right\} \tag{11}$$

In the same pattern, $\sigma_{T-1}^M(\mu)$, for $\forall \mu$, is an MPE strategy if and only if V_{T-1}^{T-1} is *achievable*. Similarly, $\{\sigma_{T-1}^M(\mu), \sigma_T^M, \rho\} \in \Gamma_{T-2}^M(\mu)$ if and only if Sender $T-1$'s payoff reaches $cl(V_{T-1}^{T-1})(\mu)$, so on and so forth. As long as $\{\sigma_1^M, \dots, \sigma_T^M, \rho^M\}$ is an MPE, this recursive process will go through until it yields the sequence of value functions $\{V_k^t\}_{\forall t, \forall k}$.

Theorem 3. *$\{\sigma_1^M, \dots, \sigma_T^M, \rho^M\}$ form an MPE if and only if, for $\forall \mu \in \Delta(\Omega)$, $k = 1, \dots, T$, $t = 1, \dots, T$ and $\sigma_t^M(\mu) = \langle \phi_1^t(\mu), \dots, \phi_m^t(\mu) \rangle_\mu$:*

$$\rho^M(\mu) \in A^*(\mu) \tag{12}$$

$$V_T^k(\mu) = E_\mu[v^k(\rho^M(\mu), \omega)] \tag{13}$$

$$cl(V_t^t)(\mu) = \sum_{j=1}^{m} p(\phi_j^t(\mu)) V_t^t(\phi_j^t(\mu)) \tag{14}$$

$$V_{t-1}^k(\mu) = \sum_{j=1}^{m} p(\phi_j^t(\mu)) V_t^k(\phi_j^t(\mu)) \tag{15}$$

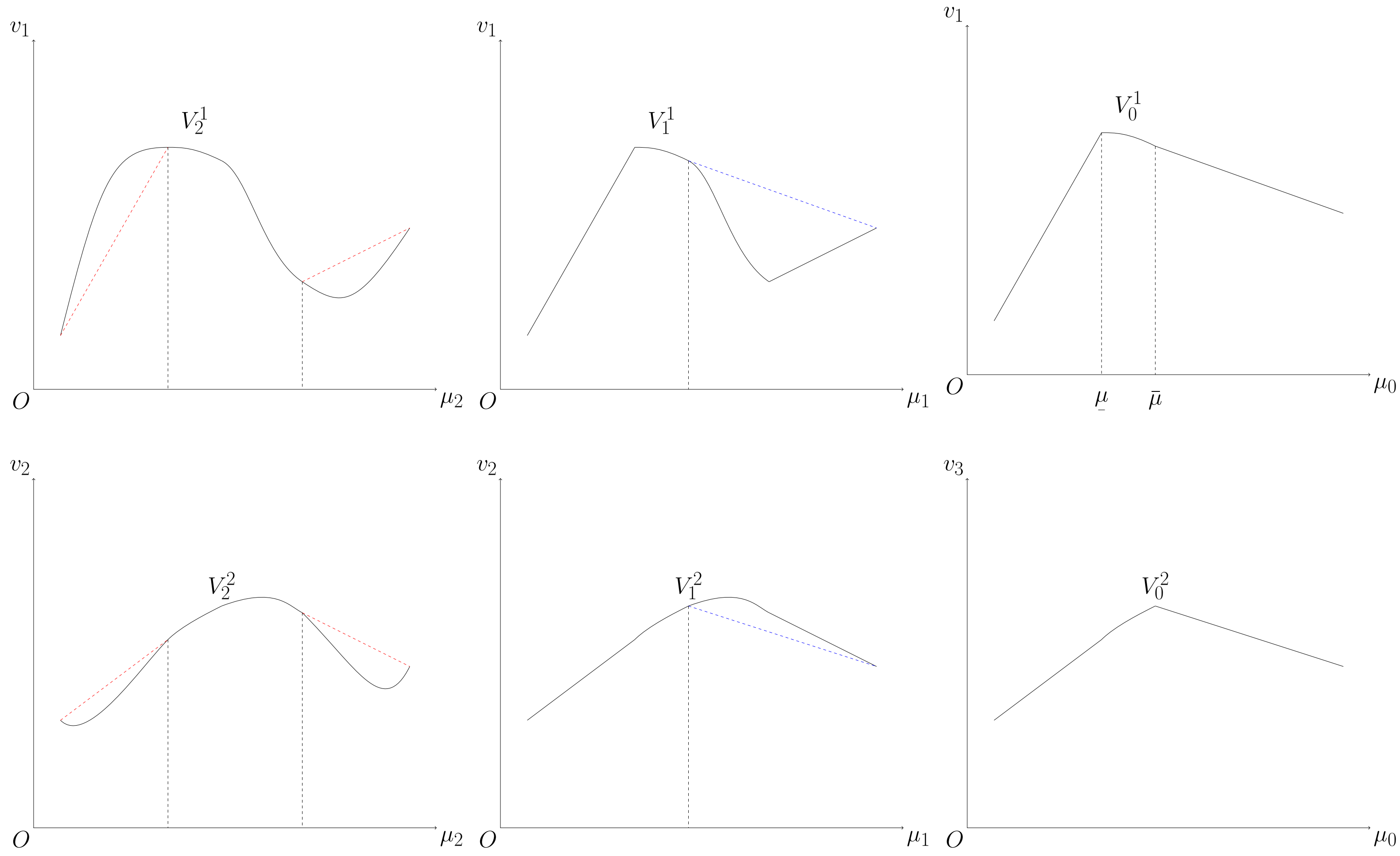


Figure 6: A General Example of an MPE

5.2 A General Example

In the MPE, the recursive concavification method takes a more transparent form. I present such an example in Figure 6. Suppose the Receiver's Markovian strategy yields senders' value functions on final beliefs in the left column. Analogous to the analysis in Section 4.2, the persuasion rents for both senders are concave closures over V_1^1 and V_2^2.[16] When Sender 1 (Sender 2) reveals information, he designs signaling rules that randomizes posteriors between the vertices of the domain of the blue (red) line the prior lies in. Also, their persuasion causes parallel randomization for the other sender that leads to transformation from V_2^1, V_2^2 to V_1^1, V_1^2, and from V_1^1, V_1^2 to V_0^1, V_0^2, respectively. Geometrically, these transformations are as obvious as connecting two value points.

Finally, the equilibrium payoffs are given as $V_0^1(\mu_0)$ and $V_0^2(\mu_0)$, and the equilibrium paths can be recovered from V_1^1 and V_2^2. If $\mu_0 \in (0, \underline{\mu})$, the equilibrium path is that Sender 1 sends $\langle 0, \underline{\mu} \rangle_{\mu_0}$, Sender 2 designs a null signaling rule; if $\mu_0 \in (\underline{\mu}, \bar{\mu})$, nobody reveals any information; if $\mu_0 \in (\bar{\mu}, 1)$, Sender 1 sends $\langle \bar{\mu}, 1 \rangle_{\mu_0}$, while Sender 2 does not reveal any information.

5.3 MPE vs. SPE

This part differentiates SPE from MPE by presenting SPE payoffs that are not in MPE. In this example, there are two states $\{\omega_0, \omega_1\}$, $\mu = P(\omega_1)$, and $\mu_0 = \frac{1}{2}$. Denote the SPE payoffs by (v_s^1, v_s^2) and the MPE payoffs by (v_m^1, v_m^2). Suppose the Receiver's action set is $[4, 6] \cup [0, 2]$ and her optimal reactions to final beliefs are represented by $A^*(\mu)$.

$$A^*(\mu) = \begin{cases} [4,6] & \text{if } \mu \in [0, \frac{1}{2}], \\ [0,2] & \text{if } \mu \in [\frac{1}{2}, 1]. \end{cases}$$

There are two senders, whose utilities are given as below.

$$v^1(a, \omega_0) = \begin{cases} 10 - a & \text{if } a \in [4,6], \\ 2 - a & \text{if } a \in [0,2]. \end{cases}$$

$$v^1(a, \omega_1) = 6 - a$$

[16] Here the concave closures of the minimum and the maximum of value correspondences coincide, Harris condition is satisfied by signaling rules that achieve the concave closure.

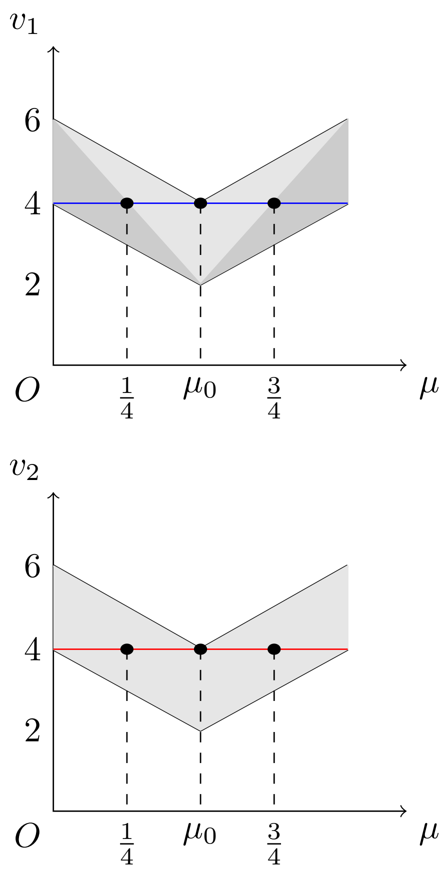


Figure 7: One SPE that is not an MPE

$$v^2(a, \omega_0) = a$$

$$v^2(a, \omega_1) = \begin{cases} a - 4 & \text{if } a \in [4, 6], \\ a + 4 & \text{if } a \in [0, 2]. \end{cases}$$

Under this payoff structure, V_2^1 and V_2^2 are depicted as the grey areas in Figure 11. One feature of this payoff structure is that v^1 and v^2 are increasing and decreasing functions in a, respectively, which means they have opposite interests in the Receiver's actions. The Receiver's "punishment" on one sender can automatically become a "reward" to the other.

For any MPE in which $v_m^2 = 4$, the Receiver must take 4 at $\mu = 0$ and 0 at $\mu = 1$, generating the lowest payoffs Sender 2 could receive in equilibrium.

Otherwise, he can deviate to the truth-telling signaling rule that brings about a payoff higher than 4. But the Receiver's behavior against Sender 2 is beneficial for Sender 1, in a sense that as long as he reveals the state, his payoff is 6, the highest possible payoff for him in this game. Therefore, $(v_m^1, v_m^2) = (6, 4)$ is the unique pair of MPE payoffs in which $v_m^2 = 4$.

In SPE, however, the Receiver has more lattitude in punishing senders for deviations from certain information structures, and consequently there will be more abundant equilibria. Focusing on a class of equilibria where Sender 2 keeps "silent" and the Receiver takes actions satisfying Harris condition of the second period, a subset of V_1^1 is shown as the deeper grey area in the upper figure.[17] It is sufficient to obtain Harris condition of the first period from this subset of V_1^1, as shown by the blue line.[18] Then, I have one SPE path: Sender 1 designs $\langle \frac{1}{4}, \frac{3}{4} \rangle_{\frac{1}{2}}$, Sender 2 sends null signals and the Receiver takes 5 at $\mu = \frac{1}{4}$ and 1 at $\mu = \frac{3}{4}$. Here, $(v_s^1, v_s^2) = (4, 4)$.

6 The Silent Equilibrium

When the equilibrium information structure is unique, in either Example 1 or Figure 6, it is an equilibrium that only Sender 1 sends a nontrivial signaling rule, while Sender 2 babbles. This is called the Silent Equilibrium.

Definition 6. *(The Silent Equilibrium) An SPE in which at most one sender reveals information on the equilibrium path.*

This section demonstrates the existence of this class of equilibria by construction.

Suppose that when the Receiver is indifferent between multiple optimal actions at a belief, she chooses those in favor of Sender T; if there are still multiple actions, among these actions she chooses those in favor of Sender $T-1$, so on and so forth. That is, the Receiver breaks the tie by benefiting senders in an order of priority: $T \succ T-1 \ldots \succ 1$.

Furthermore, I focus on *partially conservative* senders. That a sender is *conservative* means that he would rather not revealing any information if not necessary. He is *partially conservative* if he is *conservative* only when he thinks that following senders would not reveal anything; otherwise he will reveal the amount of information his followers would have revealed.

[17] Because my goal is to find a specific SPE, I do not need to characterize the whole set of V_1^1.

[18] Harris condition cannot be higher than the concave closure of the minimum of this subset, and it cannot be lower than the hyperplane going across the value points at the verticies, according to Prop. 7.

Below I formalize the main result in Theorem 4 and go through the proof in the remainder of this part. The proof is separated into several steps, some of which are postponed to the appendix for the expositional simplification.

Theorem 4. *There always exists a Silent Equilibrium.*

Proof. As below in Sec 5.3.1, 5.3.2 and 5.3.3. □

The senders' behavior creates a pattern that earlier senders prefer to reveal information his followers would have revealed, but if nobody else "speaks," they have no incentive to do so. In any equilibrium formed by these senders, Sender $T-1$ leaves Sender T nothing to "say" as he has said what Sender T would have, Sender $T-2$ makes Sender $T-1$ silent, so on and so forth. In the limit of this logic, Sender 1 is the only one who could possibly reveal any information.

6.1 Period T

Suppose the MPE of interest is $\{\sigma_1^S, \ldots, \sigma_T^S, \rho^S\}$. ρ^S is any Markovian strategy satisfying the tie breaking rule, which derives senders' value functions $\{V_T^t\}_{\forall t}$. Before diving into the recursion, I make two definitions: the *pseudo* value functions $\tilde{V}_t^t$ and the *active* set $\mathcal{G}^t$. $\tilde{V}_t^t$ is an instrumental value function treated by Sender t as his "real" value function when designing his strategy. It turns out that this *pseudo* value function is *achievable* and generate an optimal strategy σ_t^S. Next, the *active* set, $\mathcal{G}^t$, is a set of beliefs at which Sender t designs non-trivial signaling rules, prescribed by σ_t^S, to improve his welfare. In the transformation, $\mathcal{G}^t$ is an area in which $\{V_{t-1}^k\}_{\forall k}$ differs from $\{V_t^k\}_{\forall k}$. In the opposite, the complement $(\mathcal{G}^t)^c$ identifies the set of belief space where the value functions are unchanged at that step. The formal definitions of the *pseudo* value function and the *active* set are given inductively as below,

Definition 7. *The pseudo value function* $\tilde{V}_T^T = V_T^T$. *For* $t = 1, \ldots, T-1$,

$$\tilde{V}_t^t = \{ \begin{array}{ll} V_t^t & \mu \in (\mathcal{G}^{t+1})^c \\ -\infty & otherwise \end{array}$$

where $-\infty$ can be regarded as a sufficiently small constant lower than the worst payoff in the game.

Definition 8. *For* $\forall t$, *the active set* $\mathcal{G}^t = \{\mu \in \Delta(\Omega) | cl(V_t^t)(\mu) > \tilde{V}_t^t(\mu)\}$.

In period T, because the Receiver gives Sender T the highest priority, Sender T is endowed with an USC value function $V_T^T(=\tilde{V}_T^T)$. Satisfying *partial conservativeness*, Sender T's optimal strategy σ_T^S is recovered by the concavification method, such that $\sigma_T^S(\mu) := \langle \phi_1(\mu), \ldots, \phi_m(\mu)\rangle_\mu$,

$$cl(V_T^T)(\mu) = \sum_{j=1}^{m} p(\phi_j(\mu))\tilde{V}_T^T(\phi_j(\mu))$$

Based on σ_T^S, other senders' value functions $\{V_{T-1}^t\}_{\forall t}$ are generated symmetrically,

$$V_{T-1}^t(\mu) = \Big\{ \sum_{j=1}^{m} p(\phi_j(\mu))V_T^t(\phi_j(\mu)) \mid \sigma_T^S(\mu) = \langle \phi_1(\mu), \ldots, \phi_m(\mu)\rangle_\mu \Big\}$$

Here are some important lemmas for following steps.

Lemma 1. *For* $\forall \mu \in \Delta(\Omega)$ *and* $\sigma_T^S(\mu) = \langle \theta_1, \ldots, \theta_m\rangle_\mu$, $\theta_j \notin \mathcal{G}^T$, *for* $\forall j$.

Proof. If $\theta_1 \in \mathcal{G}^T$ has an optimal signaling rule $\langle \phi_1, \ldots, \phi_m\rangle_{\theta_1}$. Then $\langle\langle \phi_1, \ldots, \phi_m\rangle_{\theta_1}, \theta_2, \ldots, \theta_m\rangle_\mu$ [19] is a signaling rule that dominates $\sigma_T^S(\mu)$. □

Lemma 2. $\mathcal{G}^T$ *is open.*

Proof. In $(\mathcal{G}^T)^c$, V_T^T coincides with $cl(V_T^T)$, which is LSC. Suppose there is a sequence of beliefs $\mu^n \in (\mathcal{G}^T)^c, \forall n$ and $\{\mu^n\} \to \mu$. $\{\rho(\mu^n)\}$ is a sequence of the Receiver's optimal responses. $\rho(\mu^n) \in A^*(\mu^n), \forall n$, has a converging subsequence $\rho(\mu^{n_k}) \to a' \in A$. $a' \in A^*(\mu)$ by the Maximum Theorem. Because of the tie breaking rule, $V_T^T(\mu) \geqslant E_\mu[v^T(a', \mu)] = \lim_{k\to\infty} E_{\mu^{n_k}}[v^T(\rho(\mu^{n_k}), \mu^{n_k})] = \lim_{k\to\infty} cl(V_T^T)(\mu^{n_k}) \geqslant cl(V_T^T)(\mu)$. $\mu \in (\mathcal{G}^T)^c$, $(\mathcal{G}^T)^c$ closed. □

Lemma 3. V_T^T *is continuous in* $(\mathcal{G}^T)^c$.

Proof. See Appendix. □

In period $T-1$, Sender $T-1$ is active in face of the transformed value function V_{T-1}^{T-1}. Because Sender T is *partially conservative*, V_{T-1}^{T-1} only possibly differs from V_T^{T-1} within $\mathcal{G}^T$, where $\tilde{V}_T^T$ is defined generically as $-\infty$. Next, I take a detour to prove that V_{T-1}^{T-1} is achievable by showing the achievability of $\tilde{V}_{T-1}^{T-1}$ first.

[19] $\langle\langle \phi_1, \ldots, \phi_m\rangle_{\theta_1}, \theta_2, \ldots, \theta_m\rangle_\mu$ means that $\mu = p(\theta_1)p(\phi_1)\phi_1 + \cdots p(\theta_1)p(\phi_m)\phi_m + p(\theta_2)\theta_2 + \cdots + p(\theta_m)\theta_m$. One may be concerned with the restriction on the size of the signal space. But strictly speaking, what actually contradicts is that $\langle \phi_1, \ldots, \phi_m\rangle_{\theta_1}$ cannot generate an expected payoff that matches the concave closure, which is not restricted in the size of posteriors according to its definition in Eq. (1).

Lemma 4. $\tilde{V}_{T-1}^{T-1}$ *is USC.*

Proof. See Appendix. □

An USC $\tilde{V}_{T-1}^{T-1}$ is achievable without any doubt. V_{T-1}^{T-1}, however, can be shown to be achievable by Lemma 5, $cl(\tilde{V}_{T-1}^{T-1}) = cl(V_{T-1}^{T-1})$. One important observation used in the proof of Lemma 5 is that when $\mu_{T-1} \notin \mathcal{G}^T$, Sender T would not reveal any information as a conservative person. Therefore, $\tilde{V}_{T-1}^{T-1}(\mu) = V_{T-1}^{T-1}(\mu) = V_T^{T-1}(\mu)$, $\forall \mu \notin \mathcal{G}^T$. For $\mu \in \mathcal{G}^T$, however, it can be shown that the concave closure of $\tilde{V}_{T-1}^{T-1}$, which is supposed to lie weakly below $cl(V_{T-1}^{T-1})$, weakly dominates V_{T-1}^{T-1}. This is true only when both concave closures coincide.

Lemma 5. $cl(\tilde{V}_{T-1}^{T-1}) = cl(V_{T-1}^{T-1})$

Proof. See Appendix. □

With the knowledge of $\tilde{V}_{T-1}^{T-1}$, $\mathcal{G}^{T-1}$ is given by Definition 7. Lemma 6 shows the openness of $\mathcal{G}^{T-1}$, in parallel with Lemma 2, and that the *active* set of Sender $T-1$ contains Sender T's.

Lemma 6. $\mathcal{G}^{T-1}$ *is open and* $\mathcal{G}^{T-1} \supset \mathcal{G}^T$.

Proof. See Appendix. □

6.2 Period t, $t < T$

Lemma $1-6$ are preliminary to deriving similar results for $t < T$. Suppose Sender t designs his optimal signaling rule as if $\tilde{V}_t^t$ is his value function, which is an equivalent way of solving for the equilibrium based on similar results with Lemmas 4 and 5, that are proved in Prop. 9. Therefore, for $\forall k = 1, ..., T$, $\sigma_k^S(\mu) := \langle \phi_1, \dots, \phi_m \rangle_\mu$ is any strategy satisfying,

$$cl(V_k^k)(\mu) = \sum_{j=1}^{m} p(\phi_j) \tilde{V}_k^k(\mu) \tag{16}$$

subject to the constraint of *partial conservativeness*. Also, other senders' value functions are generated from σ_k^S,

$$V_{k-1}^t(\mu) = \left\{ \sum_{j=1}^{m} p(\phi_j(\mu)) V_k^t(\phi_j(\mu)) \mid \sigma_k^S(\mu) = \langle \phi_1(\mu), \dots, \phi_m(\mu) \rangle_\mu \right\} \tag{17}$$

$\{\sigma_1^S, \ldots, \sigma_T^S, \rho^S\}$ are constructed by the recursive concavification method as above, generating a sequence of value functions $\{V_t^k\}_{\forall t, \forall k}$. Prop. 9 is divided into two parts, the first of which includes preconditions that have been satisfied for $t = T$ by Lemmas 2 and 3. While the second part shows useful conclusions corresponding to Lemmas 1, 4, 5, 6.

Proposition 9. *(Induction) For $\forall t$, if following conditions are satisfied:*

1. *V_t^k is transformed based on Eq. (17) for $k \geq t$.*
2. *$\mathcal{G}^t$ open.*
3. *V_T^k continuous in $(\mathcal{G}^t)^c$, $\forall k \geq t$.*
4. *$\mathcal{G}^t \supset \mathcal{G}^{t+1} \cdots \supset \mathcal{G}^T$.*

It can be concluded that:

1. *$\mathcal{G}^{t-1}$ open.*
2. *V_T^k continuous in $(\mathcal{G}^{t-1})^c$, for $\forall k \geq t-1$.*
3. *$\mathcal{G}^{t-1} \supset \mathcal{G}^t$.*
4. *Let $\sigma_t^S(\mu) := \langle \phi_1(\mu), \ldots, \phi_m(\mu) \rangle_\mu$, then $\phi_j(\mu) \in (\mathcal{G}^t)^c$, for $\forall j, \forall \mu$.*
5. *$\tilde{V}_{t-1}^{t-1}$ USC.*
6. *$cl(\tilde{V}_{t-1}^{t-1}) = cl(V_{t-1}^{t-1})$.*

Proof. See Appendix. □

6.3 Proof of Theorem 3

Once the recursion proceeds to the end, I can obtain two results that are essential for the final proof. First, the *active* sets are growing as t decreases, i.e. $\mathcal{G}^1 \supset \mathcal{G}^2 \supset \cdots \supset \mathcal{G}^T$. The interpretation is that earlier senders are more active due to their prescribed personalities. Second, Sender 1's equilibrium signaling rule spreads posteriors within $(\mathcal{G}^1)^c$. Combining these two sentences, one could find that all the posteriors of Sender 1's signaling rule lie out of $\mathcal{G}^t, \forall t$. It is equivalent to say that no matter what signal is generated after Sender 1's persuasion that leads to a certain belief, no other senders would think necessary to change it. Then *conservative* senders will simply stay "silent." This argument closes the proof of Theorem 3.

7 Zero Sum Game

This section generalizes the full revelation result[20] in Example 1 to any Bayesian persuasion game with two senders who have zero-sum utilities. Consistent with Gentzkow and Kamenica (2017a) (2017b), this result has a strong economic implication that the extremist conflicts between senders would improve the information revelation to the highest level.

The intuition is that the evidence from the second sender suppresses the first sender's ambition in manipulating the Receiver's belief, such that he cannot exploit his advantage of persuasion more than telling the truth. It is unavoidable that his signals become vague once the first sender gives a strong persuasion, that tilts the Receiver to the second sender as a more reliable information source. Then the second sender can counteract the influence of the first sender by giving a slightly more convincing persuasion in the opposite direction. The first sender knows that so he reveals the true state at the beginning.

Proposition 10. *In a zero-sum sequential Bayesian persuasion model, full revelation can always be sustained as an equilibrium outcome.*

Proof. A zero-sum game implies that for any MPE, $V_t^1 + V_t^2 = c$, for $t = 0, 1, 2$, $c \in \mathbb{R}$. Let the Receiver break the tie in favor of Sender 2, then V_2^2 is USC and Sender 2's strategy is generated accordingly. The concave closure of V_2^2 is V_1^2, of which the opposite V_1^1 is convex. Most importantly, Sender 1 faces a convex value function now, and his optimal signaling rule is identified by imposing a concave closure over this convex value function, that turns out to be a hyperplane going across all vertices of the graph. It suggests that the full revelation is one optimal choice for Sender 1. Given Sender 1's choice, Sender 2's action has no effect once the true state has been disclosed. □

8 Blackwell's Order and Posterior Dispersion

In this section, I extend Blackwell's order to a more specific case in which more informative signaling rules are equivalent to having more dispersed posteriors under a certain condition. By "dispersed" I mean that the convex hull of the induced posteriors contains that of a less informative signaling rule.

[20]The full revealing equilibrium is a desirable outcome that attracts attention from many related work. Krishna and Morgan (2001), Battaglini (2002), and Ambrus and Takahashi (2008) discuss conditions under which full revelation would occur in a multi-sender cheap talk. Milgrom and Roberts (1986) study when sophistication of the decision maker and the competition among senders are sufficient for this result.

This geometric way of expressing Blackwell's order is useful for the analysis in Sections 9 and 10. The notation is inherited from Section 3.3.3. Specifically, $\pi = \langle \lambda_1, \dots, \lambda_{m_1} \rangle_{\mu_0}$ and $\pi' = \langle \mu_1, \dots, \mu_{m_2} \rangle_{\mu_0}$, i.e. $\mu_0 = p_1\lambda_1 + \cdots + p_{m_1}\lambda_{m_1}$ and $\mu_0 = q_1\mu_1 + \cdots + q_{m_2}\mu_{m_2}$.

Initially, I propose the definition of *dominance* to capture the idea of dispersed posteriors.

Definition 9. *(Dominance)* π' dominates π *if for* $\forall j$, $\lambda_j \in conv\{\mu_1, \dots, \mu_{m_2}\}$, *denoted by* $\pi' > \pi$.

Next, I define that a signaling rule π' is *singular* as below.

Definition 10. *(Singularity) A signaling rule π is singular when the number of its distinct posteriors*[21] *is equal to the dimension of their convex hull plus 1.*

With the property of *singularity*, the signaling rule's convex combination coefficients for any point in the convex hull of its posteriors are unique.[22] Then I can extend Blackwell's theorem as below.

Theorem 5. *In a game with a common prior, if π' is singular, $\pi' \succ \pi$ iff $\pi' \supset \pi$ iff $\pi' > \pi$.*

Proof. See Appendix. □

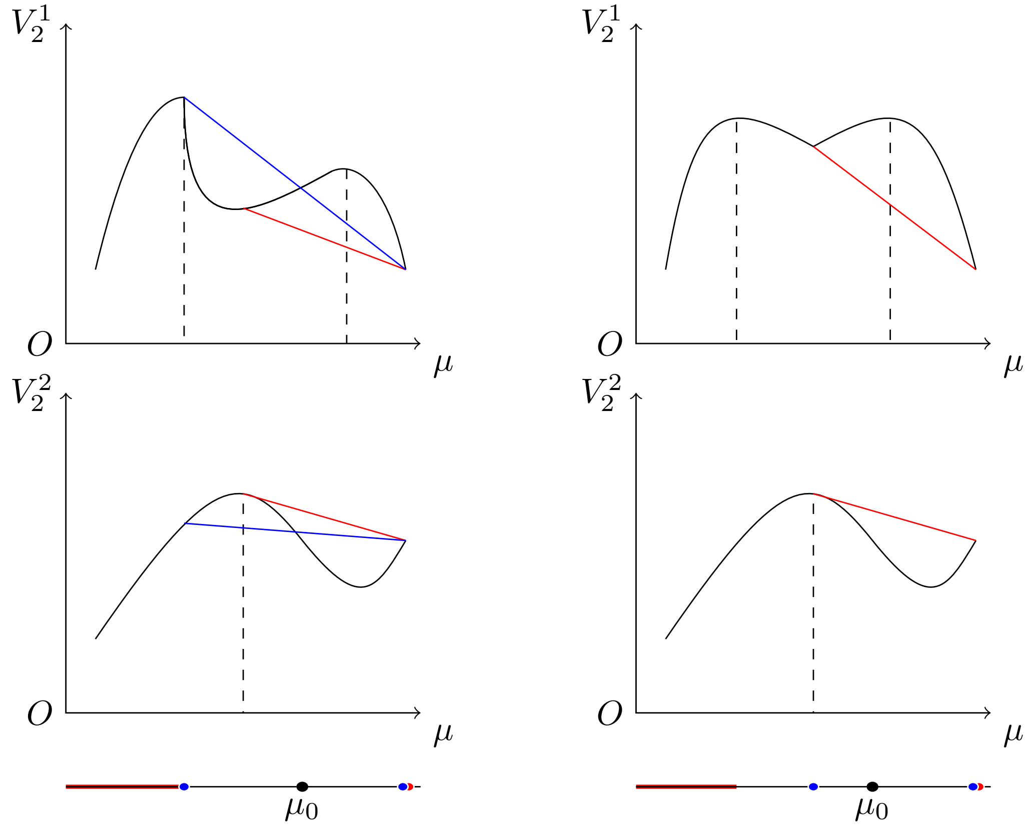
This geometric version of Blackwell's theorem derives from the property of *Bayesian plausibility* that links distributions of posteriors with signaling rules. Absent the common prior, the traditional Blackwell ordering can not be expressed this way.

Though Theorem 5 is restrictive, it is sufficient for the analysis in Sections 9 and 10. The reason is that I only focus on a specific case with unique equilibrium, where the equilibrium is represented by a *singular* signaling rule. Uniqueness in this paper means that the distribution of the induced posteriors and the associated optimal actions are unique.

Proposition 11. *For a Bayesian persuasion game that has a unique equilibrium, this equilibrium is a Silent Equilibrium, in which one sender sends a singular signaling rule.*

[21]A posterior induced by different signals counts one. Under Blackwell's order, it is equivalent to view signals inducing the same posterior as one signal.

[22]Fixed any posterior, vectors pointing from this posterior to any other posteriors are linearly independent and form a basis for their convex hull. Therefore, any point within this convex hull has unique coordinate values.



(a) Comparison 1

(b) Comparison 2

Figure 8: Comparison with Simultaneous Games

Proof. Because of Theorem 4, the unique equilibrium can be a Silent Equilibrium. WLOG, let Sender 1 designs a signaling rule $\langle \phi_1, \ldots, \phi_m \rangle_{\mu_0}$ such that $\{\phi_j\}_{\forall j}$ are distinct with each other. The optimality of the signaling rule requires that $\{V_1^1(\phi_i)\}_{i=1}^m$ are on a hyperplane $\mathcal{H}$ that weakly dominates $cl(V_1^1)$.

Suppose $dim(conv\{\phi_1, \ldots, \phi_m\}) = n$ and $m > n+1$. By Carathéodory's Theorem, there exist $\phi_{k_1}, \ldots, \phi_{k_{n+1}}$, $k_i \in \{1, \ldots, m\}, \forall i$, such that $\mu_0 = b_1\phi_{k_1} + \cdots + b_{n+1}\phi_{k_{n+1}}$, $b_i \in [0,1], \forall j, \sum_j b_j = 1$. Because $V_1^1(\phi_{k_1}), \ldots, V_1^1(\phi_{k_{n+1}})$ are on $\mathcal{H}$, Therefore, $\langle \phi_{k_1}, \ldots, \phi_{k_{n+1}} \rangle_{\mu_0}$ is another optimal signaling rule for Sender 1 as an equilibrium outcome, contradiction to uniqueness. □

9 Comparison with the Coordinated Simultaneous Bayesian Persuasion Model

Gentzkow and Kamenica (2017b) study a multi-sender persuasion game that shares the coordinated information structure but assumes simultaneity in senders' move. Their equilibrium predictions depart from this paper dramatically, which is illustrated by two examples in this section.

They characterize the equilibrium outcome, i.e. the information structure in equilibrium, by emphasizing the intersection of all *unimprovable* sets,[23] as the feasible location of posteriors. For example, in Figure 8, the intersections of *unimprovable* sets are denoted by the red area; the equilibrium outcome is the set of signaling rules with posteriors lying in this area.

In Figure 8, I present simultaneous and sequential equilibrium outcomes with red and blue colors, respectively. Hereafter, I use red segments as the value transformation of period 2 and blue segments as that of period 1. The dashed lines point out boundaries of *unimprovable* sets. In subfigures (a) and (b), equilibria of a sequential game are characterized by the two blue points as posteriors. Compared to the simultaneous games, the sequential model narrows down the set of equilibria. However, it is not true that the former is always a refinement of the latter. In (a), the equilibrium outcome of the sequential game belongs to that of the simultaneous game; but in (b), the sequential equilibrium is strictly less informative than any simultaneous equilibrium based on Theorem 5.

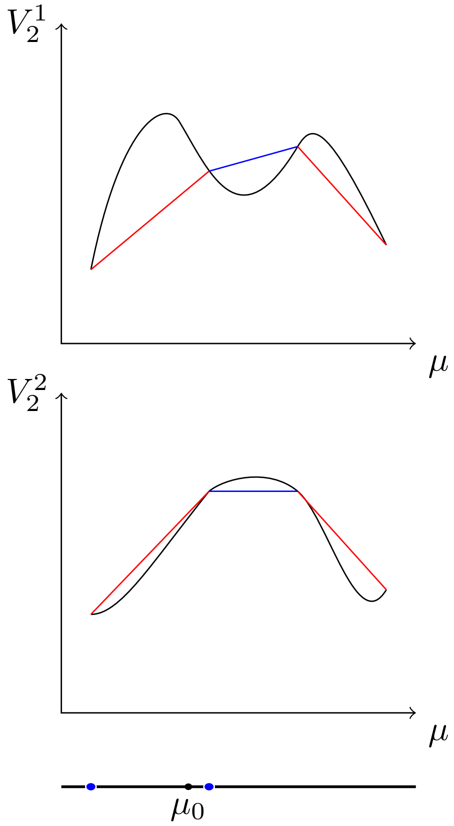
A stronger result, given unique equilibrium, is pointed out by Li and Norman (2017b), as claimed in Prop. 12. Their proof draws on the fact that the equilibrium outcome of a sequential case is a "one-step" equilibrium (Silent Equilibrium), which also holds in this paper (Theorem 4). Moreover, the two examples in this section support this result.

Proposition 12. *(Li and Norman 2017b) If a coordinated sequential Bayesian persuasion model has a unique equilibrium outcome, its corresponding simultaneous model can not have a less informative equilibrium outcome.*

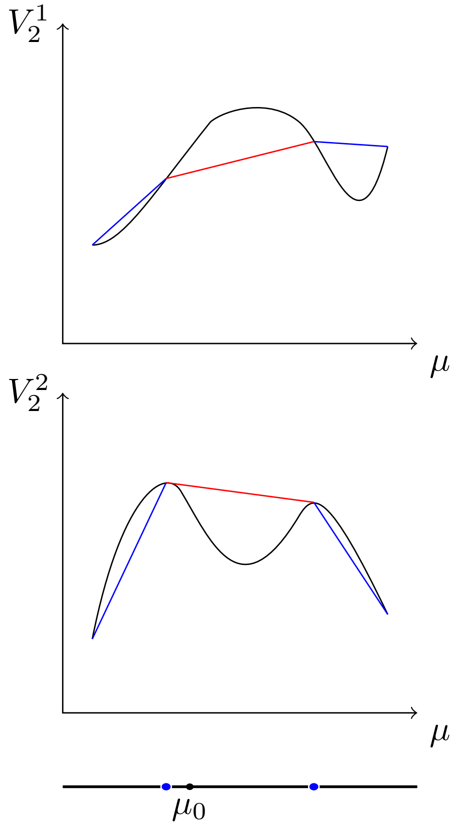
Proof. Suppose an information structure π is less informative than the Silent Equilibrium. If π is an equilibrium outcome in the simultaneous game, Sender 1 knows that if he chooses π no following senders have incentive to reveal information.[24] Apparently, Sender 1 prefers the Silent Equilibrium more than π as the Silent Equilibrium is the unique equilibrium in the sequential game.

[23]The area of belief space where the concave closure coincides with the value function.

[24]Given a "prior" in the *unimprovable* set, it is optimal for any sender to send a null signal, which form an equilibrium.



(a) Order 1



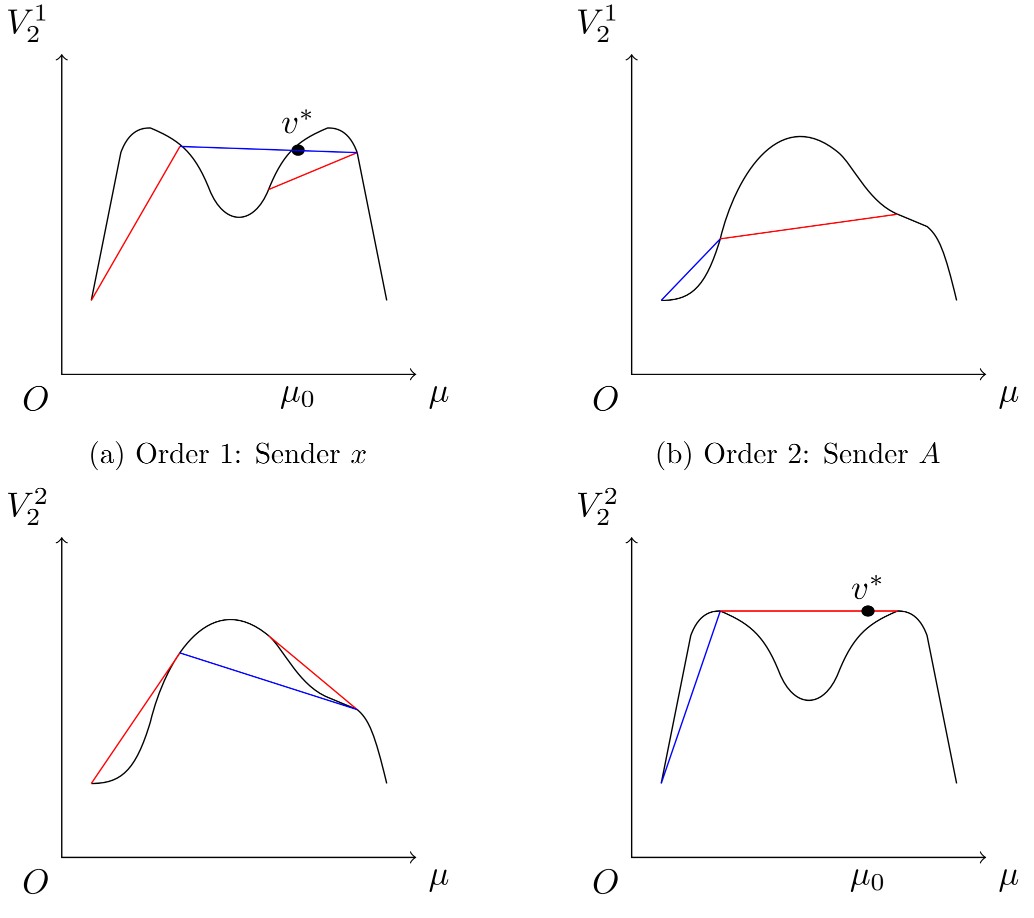
(b) Order 2

Figure 9: Equilibrium outcomes in reversed orders

It conversely implies that Sender 1 can deviate to the Silent Equilibrium to strictly improve his payoff in the simultaneous model, which rejects π as an equilibrium in the simultaneous game, contradiction. $\square$

10 Order of Persuasion

So far I have been assuming fixed order of senders' move, while in this section I make comparative static analysis by relaxing this assumption. Even for the same set of senders and the Receiver, interchanging their persuasion order could significantly affect senders' behavior. All senders are exposed to the risk of excessive information revelation from their followers, however, the direction of this effect is ambiguous. In some cases, this threat could deter



(a) Order 1: Sender x

(b) Order 2: Sender A

(c) Order 1: Sender A

(d) Order 2: Sender x

Figure 10: Sender x plays with Sender A

earlier senders from “saying” anything (Figure 9, Order 2);[25] but it is also possible that this threat would stimulate information revelation (Figure 10, Order 2). Also, the impact on informativeness is obscure. In Figure 9, two reversed orders result in equilibria with two uncomparable signaling rules according to Theorem 5.

Another interesting question is about a sender’s favorite moving spot. One may conjecture that for any sender, it is the best for him to move as late as possible, because then he has the later word power to adjust his revelation flexibly. But an earlier mover has an advantage of setting up the minimal level of revelation, which could outweigh the benefits from late word power.

[25] There is an equilibrium where the first sender keeps “silent,” which is not the case when he is the single sender.

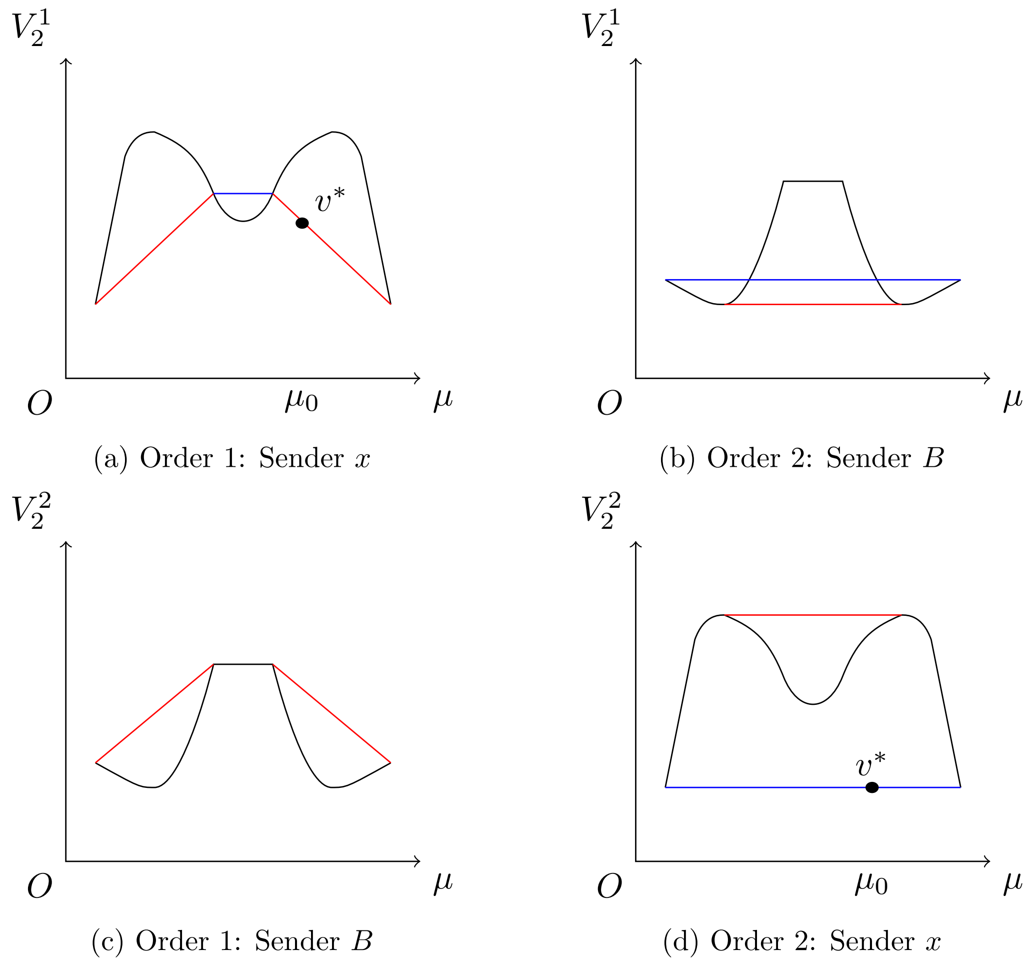


Figure 11: Sender x plays with Sender B

It turns out that even one specific sender could have different preferences over moving spots depending on his opponent. To justify, I construct two examples where Sender x plays with Sender A and Sender B, separately. The results are illustrated in Figure 10 and 11, where Sender x's equilibrium payoffs are denoted by v^*. As the pictures show, when playing with Sender A, Sender x would like to move afterwards; but when his opponent becomes Sender B, he prefers to move first.

These examples emphasize the sensitivity of equilibrium outcomes to the order of persuasion, and suggest the necessity of examining them case by case. Generally speaking, because the permutation of order is finite, one can solve the whole set of games with different orders and then make comparisons.

11 Conclusion

This paper extends the Bayesian persuasion model (Kamenica and Gentzkow 2011) into a multi-sender sequential case where subsequent players are able to observe previous signaling rules and their realizations. This setting allows senders to correlate probability of their signals with previous senders', and to accurately pick a weakly more transparent information structure than the existing one. This class of games, called "coordinated sequential Bayesian persuasion," gives rise to an environment different from the case in which senders move simultaneously (Gentzkow and Kamenica 2017b).

The main contributions contain two aspects. First, this paper solves for SPE of a Bayesian persuasion model. SPE is more general an equilibrium concept than those in the literature, which exhausts the strategic information release and identifies the ranges of equilibrium payoffs to senders. Second, in the spirit of Harris (1985), I develop the recursive concavification method to characterize the SPE paths, that draws on decentralization of the overall persuasion phase into single-sender Bayesian persuasion games.

Another implication of the coordinated information environment is the existence of a specific type of equilibria, the Silent Equilibrium. Under certain tie breaking rules, senders, who have flexibility in increasing the information revelation level, would reveal what subsequent senders would have revealed, so that it is only necessary for one sender to design a non-trivial signaling rule.

Also, this paper proves that full revelation can always be maintained as an, in many cases unique, equilibrium outcome in a zero sum game. This result shows the improvement caused by competition on information revelation in the persuasion game. An intuitive example consisting of the judge, prosecutor and attorney is proposed at the beginning for illustration.

Finally, this paper extends Blackwell's theorem in a game with a common prior. It turns out that Blackwell's order can be conditionally represented by the posterior dispersion of an information structure. This geometric version of Blackwell's theorem provides a convenient way of examing informativeness of equilibrium outcomes in certain examples, by which I compare simultaneous with sequential games and conclude the sensitivity of equilibrium outcomes to the order of persuasion.

References

[1] Ambrus, Attila and Satoru Takahashi [2008], "Multi-sender cheap talk with restricted state spaces," *Theoretical Economics* **3.1**,

1-27.

[2] Aumann, Robert J., Michael Maschler, and Richard E. Stearns [1995], "Repeated games with incomplete information," MIT Press.

[3] Battaglini, Marco [2002], "Multiple referrals and multidimensional cheap talk," *Econometrica* **70.4**, 1379-1401.

[4] Bergemann, Dirk and Stephen Morris [2016] "Bayes correlated equilibrium and the comparison of information structures in games," *Theoretical Economics* **11.2**, 487-522.

[5] Blackwell, David [1951], "Comparison of experiments," Proceedings of the second Berkeley symposium on mathematical statistics and probability. The Regents of the University of California.

[6] Blackwell, David [1953], "Equivalent comparisons of experiments," *The annals of mathematical statistics* **24.2**, 265-272.

[7] Crawford, Vincent P and Joel Sobel [1982], "Strategic information transmission," *Econometrica: Journal of the Econometric Society* 1431-1451.

[8] Ely, Jeffrey C [2017] "Beeps," *The American Economic Review* **107.1**, 31-53.

[9] Ewald, Günter [2012] "Combinatorial convexity and algebraic geometry," Springer Science & Business Media, **168**.

[10] Forges, Françoise and Koessler, Frédéric [2008], "Long persuasion games," *Journal of Economic Theory* **143.1**, 1-35.

[11] Gentzkow, Matthew, and Emir Kamenica [2017a], "Competition in persuasion," *The Review of Economic Studies*, **84.1**, 300-322.

[12] Gentzkow, Matthew, and Emir Kamenica [2017b], "Bayesian persuasion with multiple senders and rich signal spaces," *Games and Economic Behavior*, **104**, 411-429.

[13] Grossman, Sanford J [1981], "The informational role of warranties and private disclosure about product quality," *The Journal of Law and Economics*, **24.3** 461-483.

[14] Harris, Christopher [1985], "Existence and characterization of perfect equilibrium in games of perfect information." *Econometrica: Journal of the Econometric Society*, 613-628.

[15] Harris, Christopher, Philip Reny, and Arthur Robson [1995], "The existence of subgame-perfect equilibrium in continuous games with almost perfect information: A case for public randomization." *Econometrica: Journal of the Econometric Society*, 507-544.

[16] Hellwig, Martin, et al [1990], "Subgame perfect equilibrium in continuous games of perfect information: an elementary approach to existence and approximation by discrete games." *Journal of Economic Theory*, **52.2**, 406-422.

[17] Hiriart-Urruty, Jean-Baptiste, and Claude Lemaréchal [2012], "Fundamentals of convex analysis." Springer Science and Business Media.

[18] Kamenica, Emir, and Matthew Gentzkow [2011], "Bayesian Persuasion," *The American Economic Review*, **101.6**, 2590-2615.

[19] Krishna, Vijay and John Morgan [2001], "A model of expertise," *The Quarterly Journal of Economics*, **116.2**, 747-775.

[20] Li, Fei, and Peter Norman [2017a] "On Bayesian persuasion with multiple senders," Working paper.

[21] Li, Fei, and Peter Norman [2017b] "Sequential Persuasion," Working paper.

[22] Milgrom, Paul R [1981], "Good news and bad news: Representation theorems and applications," *The Bell Journal of Economics* 380-391.

[23] Milgrom, Paul and John Roberts [1986], "Relying on the information of interested parties," *The RAND Journal of Economics* 18-32.

12 Appendix

12.1 Proposition 1

Proof. WLOG, it is sufficient to show that the proposition holds in a two-sender case. Suppose there are two senders, Sender 1 and Sender 2, and Sender 2 moves later. In Case 1 both senders propose coordinated signaling rules $\{\pi_1^c, \pi_2^c\}$, while in Case 2 they propose a constellation of independent signaling rules, $\{\pi_1, \{\pi_{2,s}\}_{s\in S_1}\}$ where Sender 2 varies his signaling rules over different signals received. In this part, I use $P_C(\cdot)$ to represent the probability of the coordinated signaling rules and $P_I(\cdot)$ that of the independent ones. s_k^i represents the ith signal sent by Sender k.

(Case 1 $\Rightarrow$ Case 2)

Suppose $\{\pi_1^c, \pi_2^c\}$ is given. Construct the independent signaling rule π_1 as that $P_I(s_1^i|\omega) = P_C(s_1^i|\omega)$ and $\{\pi_{2,s}\}_{S_1}$ as that $P_I(s_2^j|s_1^i, \omega) = P_C(s_1^i, s_2^j|\omega)/P_C(s_1^i|\omega)$, where $\sum_{j=1}^m P_I(s_2^j|s_1^i, \omega) = 1$ for $\forall i$, $\forall j$, $\forall\omega$. By this construction, for $\forall i$, $\forall j$, $\forall\omega$,

$$\begin{aligned} P_I(s_1^i, s_2^j|\omega) &= P_I(s_2^j|s_1^i, \omega)P_I(s_1^i|\omega) \\ &= P_I(s_2^j|s_1^i, \omega)P_C(s_1^i|\omega) \\ &= P_C(s_1^i, s_2^j|\omega) \end{aligned}$$

(Case 2 $\Rightarrow$ Case 1)

Now $\{\pi_1, \{\pi_{2,s}\}_{s\in S_1}\}$ are given. Let π_1^c satisfy that $P_C(s_1^i|\omega) = P_I(s_1^i|\omega)$ and π_2^c satisfy that $P_C(s_1^i, s_2^j|\omega) = P_I(s_2^j|s_1^i, \omega)P_I(s_1^i|\omega)$ for $\forall i, \forall j, \forall\omega$. To verify that π_2^c is consistent with π_1^c, we have that $\sum_{j=1}^m P_C(s_1^i, s_2^j|\omega) = \sum_{j=1}^m P_I(s_2^j|s_1^i, \omega)$ $P_I(s_1^i|\omega) = P_I(s_1^i|\omega) = P_C(s_1^i|\omega)$ for $\forall i, \forall\omega$. Furthermore, $\{\pi_1^c, \pi_2^c\}$ generates the same information structure because $P_C(s_1^i, s_2^j|\omega) = P_I(s_2^j|s_1^i, \omega)P_I(s_1^i|\omega) = P_I(s_1^i, s_2^j|\omega)$, for $\forall i, \forall j, \forall\omega$. □

12.2 Proposition 5

Proof. This proof is divided into 3 parts. I devote the first two parts to showing that $\{V_{T-1}^t\}_{\forall t}$ have close graphs, which provides the general idea of the whole proof. Later on, the closedness of $\{V_k^t\}_{\forall t,\forall k}$ is concluded by induction.

Step 1: V_{T-1}^T has a closed graph.

For $\forall(\theta^n, v^n)$, $v^n \in V_{T-1}^T(\theta^n)$, i.e. $\exists\langle\phi_1^n, \ldots, \phi_m^n\rangle_{\theta^n}$, such that $\sum_{j=1}^m p(\phi_j^n)w_j^n = v^n$, $w_j^n \in V_T^T(\phi_j^n)$ and $\sum^m p(\phi_j^n)w_j^n \geqslant cl(\underline{V}_T^T)(\theta^n)$.

If $\{\theta^n, v^n\} \to (\theta^*, v^*)$, the corresponding signaling rules of Sender T associated with $\{\theta^n, v^n\}$ are $\{\pi_T^n\} \in \prod^l \Delta(S)$. Because the space of signaling rules is compact, there exists a subsequence $\{\pi_T^{n_k}\} \to \pi_T^* \in \prod^l \Delta(S)$.

Lemma 7. *Suppose* π_T^* *results in* $\langle \phi_1^*, \ldots, \phi_m^* \rangle_{\theta^*}$*, then* $\langle \phi_1^{n_k}, \ldots, \phi_m^{n_k} \rangle_{\theta^{n_k}} \to \langle \phi_1^*, \ldots, \phi_m^* \rangle_{\theta^*}$.

Proof. To show the convergence, I start from the convergence of coefficients and then the posteriors. The convergence of coefficients is directly given by:

$$p(\phi_j^{n_k}) = \sum_{\omega \in \Omega} \theta^{n_k}(\omega)\pi_T^{n_k}(s^j|\omega) \to \sum_{\omega \in \Omega} \theta^*(\omega)\pi_T^*(s^j|\omega) = p(\phi_j^*)$$

for $\forall j = 1, \ldots, m$. Then discuss the convergence of posteriors in two cases: 1) $p(\phi_j^*) \neq 0$. $\phi_j^{n_k}(\omega) = \theta^{n_k}(\omega)\pi_T^{n_k}(s^j|\omega)/p(\phi_j^{n_k}) \to \phi^*(\omega)$, for $\forall \omega \in \Omega$; 2) $p(\phi_j^*) = 0$. Select the subsequence such that $\{\phi_j^{n_k}\}$ converges to ϕ_j^*, and because signal j will disappear, ϕ_j^* can be assigned as the limit of $\{\phi_j^{n_k}\}$.

□

Because V_T^T is closed and A compact, the subsequence $\{n_k\}$ can be further selected such that the actions supporting $\{w_j^{n_k}\}$, i.e. $\{a^{n_k}\}$, converge to $\bar{a}_j \in A_j^*(\phi_j^*)$, for $\forall j$. Due to continuous payoff functions, the values from subgames, $w_j^{n_k} \to w_j^*$, for $\forall j$. The existence of such a subsequence is justified by the Weierstrass Theorem, based on finite action profiles on equilibrium paths.

Lemma 8. $(\pi_T^*, \{\bar{a}_j\}_{s^j \in S_T}) \in \bar{\Gamma}_{T-1}(\theta^*)$

Because a concave function is LSC over $\Delta(\Omega)$ (continuous in $int(\Delta(\Omega))$).

$$\sum_{j=1}^{m} p(\phi_j^*)w_j^* = \lim_{k \to \infty} \sum_{j=1}^{m} p(\phi_j^{n_k})w_j^{n_k} \geqslant \lim_{k \to \infty} cl(\underline{V}_T^T)(\theta^{n_k}) \geqslant cl(\underline{V}_T^T)(\theta^*) \quad (18)$$

Then Prop. 3 concludes Lemma 2.

Suppose the equilibrium payoff of $(\pi_T^*, \{\bar{a}_j\}_{s^j \in S_T})$ for Sender T is v. Then $v \in V_{T-1}^T(\theta^*)$, and because v^T is continuous and $\pi_T^{n_k} \to \pi_T^*$,

$$
\begin{aligned}
v &= \sum_{\omega\in\Omega}\sum_{s_j\in S_T} v^T(\bar{a}_j,\omega)\cdot\pi_T^*(s^j|\omega)\cdot\mu_{T-1}(\omega) \\
&= \lim_{k\to\infty}\sum_{\omega\in\Omega}\sum_{s^j\in S_T} v^T(a_j^{n_k},\omega)\cdot\pi_T^{n_k}(s^j|\omega)\cdot\mu_{T-1}(\omega) \\
&= \lim_{n\to\infty} v^n \\
&= v^*
\end{aligned}
$$

Therefore $v^* \in V_{T-1}^T(\theta^*)$.

Step 2: $\{V_{T-1}^t\}_{\forall t}$ have closed graphs.

For any $(\theta^n, \bar{v}^n)$ such that $\bar{v}^n \in V_{T-1}^t(\theta^n)$ and $(\theta^n, \bar{v}^n) \to (\theta, \bar{v})$, there is corresponding $(\theta^n, v^n) \subset graph(V_{T-1}^T)$ supported by the same SPE. Because $graph(V_{T-1}^T)$ is compact, $\exists$ a subsequence $\{(\theta^{n_k}, v^{n_k})\} \to (\theta^*, v^*) \in graph(V_{T-1}^T)$.

From above, the subsequence $\{(\theta^{n_k}, v^{n_k})\}$ can also be selected such that they represent a converging sequence of SPE paths.

$$
\{\pi^{n_k}, \{a_j^{n_k}\}_{s^j\in S_T}\} \to \{\pi_T^*, \{a_j^*\}_{s^j\in S_T}\}
$$

Additionally, v^t is continuous, so

$$
\begin{aligned}
\bar{v} &= \lim_{n\to\infty} \bar{v}^n \\
&= \lim_{k\to\infty}\sum_{\omega\in\Omega}\sum_{s^j\in S_T} v^t(a_j^{n_k},\omega)\pi_T^{n_k}(s^j|\omega)\mu_{T-1}(\omega) \\
&= \sum_{\omega\in\Omega}\sum_{s^j\in S_T} v^t(a_j^*,\omega)\pi_T^*(s^j|\omega)\mu_{T-1}(\omega) \\
&= v^*
\end{aligned}
$$

where v^* is Sender t's value of $(\pi_T^*, \{a_j^*\}_{s^j\in S_T})$. □

Step 3: $\{V_t^k\}_{\forall t,\forall k}$ have closed graphs.

Lemma 9. *Suppose for* $t, 1 \leqslant t \leqslant T$, V_t^k *is non-empty and has a closed graph for* $\forall k = 1, \ldots, T$. *And for any sequence* $(\phi^n, w^n) \to (\phi, w) \in graph(V_t^k)$, *we have in accordance a subsequence of SPE paths $\bar{\gamma}(\phi^{n_i})$ that yield payoffs w^{n_i} for Sender k, converging to a SPE path $\bar{\gamma}(\phi)$ that yields payoff w for Sender k. Then, V^k_{t-1} is non-empty and closed for $\forall k$, and for any sequence $(\phi'^n, w'^n) \to (\phi', w') \in graph(V^k_{t-1})$, there is in accordence a subsequence of SPE paths $\bar{\gamma}(\phi'^{n_i})$ that yields payoff w'^{n_i}, converging to a SPE path $\bar{\gamma}(\phi')$ that yields payoff w' for Sender k.*

Proof. (non-emptiness) The proof of non-emptiness of $V^k_{t-1}, k = 1, \ldots, T$ could be given in the same way as that in Theorem 2, based on the fact that V^t_t is closed.

(Closedness) $\forall(\theta^n, v^n)$ such that $v^n \in V^t_{t-1}(\theta^n)$ i.e. $\exists\{\langle\phi^n_1, \ldots, \phi^n_m\rangle_{\theta^n}\}$ such that $\sum_{j=1}^m p(\phi^n_j)w^n_j = v^n$, $w^n_j \in V^k_k(\phi^n_j)$ and $v^n \geq cl(\underset{-}{V}^t_t)(\theta^*)$.

Suppose $(\theta^n, v^n) \to (\theta^*, v^*)$ and their corresponding signaling rules $\{\pi^n_t\}$ yield $\{\langle\phi^n_1, \ldots, \phi^n_m\rangle_{\theta^n}\}$. By assumptions, there is a subsequence $\{n_i\}_{i=1}^\infty$ such that $\{\pi^{n_i}_t\} \to \{\pi^*_t\}$, $w^{n_i}_j \to w^*_j, \forall j = 1, \ldots, m$ and $\bar{\gamma}^{n_i}_j(\phi^{n_i}_j) \to \bar{\gamma}^*_j, \forall j = 1, \ldots, m$. Here $\bar{\gamma}^*_j$ simply denotes the SPE path starting from the point when Sender k sends signal j.

$(\pi^*_t, \{\bar{\gamma}^*_j\}_{j=1}^m)$ is a SPE path according to Theorem 1 because,

$$\sum_{j=1}^m p(\phi^*_j)w^*_j = \lim_{i\to\infty}\sum_{j=1}^m p(\phi^{n_i}_j)w^{n_i}_j \geqslant \lim_{i\to\infty} cl(\underset{-}{V}^t_t)(\theta^{n_i}) \geqslant cl(\underset{-}{V}^t_t)(\theta^*)$$

Relying on the continuity of v^t and the convergence of SPE paths and prior beliefs,

$$\begin{aligned}
v^* &= \lim_{n\to\infty} v^n \\
&= \lim_{i\to\infty}\sum_{\omega\in\Omega}\sum_{s\in S_t\times\cdots\times S_T} v^t(a^{n_i}(s), \omega)\pi^{n_i}_t(s_t|\omega)\cdots\pi^{n_i}_T(s_T|\omega)\theta^n(\omega) \\
&= \sum_{\omega\in\Omega}\sum_{s\in S_t\times\cdots\times S_T} v^t(a^*(s), \omega)\pi^*_t(s_t|\omega)\cdots\pi^*_T(s_T|\omega)\theta^*(\omega)
\end{aligned}$$

Therefore, $v^* \in V^t_{t-1}(\theta^*)$. For $V^k_{t-1}, k \neq t$, we can adopt the same step as in the proof of *Step 2*. For $\forall(\bar{\theta}^n, \bar{v}^n) \subset graph(V^t_{k-1}) \to (\bar{\theta}, \bar{v})$, there is a convergent sequence of subgame paths that, given continuity of v^k, the convergence of prior beliefs and SPE paths implies that $\bar{v} \in V^k_{t-1}(\bar{\theta})$. □

12.3 Proposition 6

Proof. It suffices to demonstrate the equivalence between $\{V^k_t\}_{\forall k,\forall t}$ in both games. $\{V^k_T\}_{\forall k}$ are dependent on the Receiver's actions and therefore in-

variant in both games. Inductively, if $\{V^k_{t+1}\}_{\forall k}$ are identical, Harris condition $cl(\underline{V}^{t+1}_{t+1})$ is the same for both games. Because the transformation from $\{V^k_{t+1}\}_{\forall k}$ is no more than taking convex combinations of points in the graph of $\{V^k_{t+1}\}_{\forall k}$ subject to Harris condition, $\{V^k_t\}_{\forall k}$ are the same in Γ_0 and Γ_∞ according to the Carathéodory's Theorem, taking into account that $graph(V^1_{t+1}, \dots, V^T_{t+1})$
$\subset \mathbb{R}^{T+l-1}$, for $\forall t$. □

12.4 Theorem 4

Lemma 3

Proof. In $(\mathcal{G}^T)^c$, V^T_T coincides with $cl(V^T_T)$, which is continuous in $int(\Delta(\Omega))$ and LSC over the whole belief space $\Delta(\Omega)$. Therefore, I only need to check the boundary of the belief space.

For any $\mu \in bd(\Delta(\Omega))$ and $\{\mu^n\} \subset int(\Delta(\Omega))$ such that $\{\mu, \{\mu^n\}\} \subset (\mathcal{G}^T)^c$ and $\{\mu^n\} \to \mu$, I want to show that $V^T_T(\mu) = \lim_{n\to\infty} V^T_T(\mu^n)$.

Because V^T_T is LSC in $(\mathcal{G}^T)^c$, $V^T_T(\mu) \leq \underline{\lim}_{n\to\infty} V^T_T(\mu^n)$. It only remains to be shown that $V^T_T(\mu) \geq \overline{\lim}_{n\to\infty} V^T_T(\mu^n)$. Because A is compact, there exists a subsequence $\{\mu^{n_k}\}$ s.t. $\{\rho(\mu^{n_k})\} \to a'$. By the Maximum Theorem, $a' \in A^*(\mu)$. Because the Receiver favors Sender T in indifference,

$$V^T_T(\mu) \geq E_\mu[v^T(a', \mu)] = \lim_{k\to\infty} E_{\mu^{n_k}}[v^T(\rho(\mu^{n_k}), \mu^{n_k})] = \lim_{k\to\infty} V^T_T(\mu^{n_k})$$

For any $\mu \in bd(\Delta(\Omega))$, $\{\mu^n\} \subset bd(\Delta(\Omega)))$ such that $\{\mu, \{\mu^n\}\} \subset (\mathcal{G}^T)^c$ and $\{\mu^n\} \to \mu$. The above analysis applies.

□

Lemma 4

Proof. For $\forall \mu \in \mathcal{G}^T$, because $\mathcal{G}^T$ is open, it is trivial $\tilde{V}^{T-1}_{T-1}$ USC at μ.

For $\forall \mu \in (\mathcal{G}^T)^c$, if $\{\mu^n\} \to \mu$ from $\mathcal{G}^T$, the statement is trivially true. Otherwise, if $\{\mu^n\} \to \mu$ from $(\mathcal{G}^T)^c$ and suppose $\overline{\lim}_{n\to\infty} \tilde{V}^{T-1}_{T-1}(\mu^n) > \tilde{V}^{T-1}_{T-1}(\mu)$, which implies $\overline{\lim}_{n\to\infty} V^{T-1}_{T-1}(\mu^n) > V^{T-1}_{T-1}(\mu)$. There exists a subsequence $\{n_k\}$ such that $\lim_{k\to\infty} V^{T-1}_{T-1}(\mu^{n_k}) = \overline{\lim}_{n\to\infty} V^{T-1}_{T-1}(\mu)$ and $\rho(\mu^{n_k}) \to a' \in A^*(\mu)$. By the Maximum Theorem, $E_{\mu^{n_k}}[v^T(\rho(\mu^{n_k}), \omega)] = V^T_T(\mu^{n_k})$ implies $E_\mu[v^T(a', \mu)] = V^T_T(\mu)$. So a' is the optimal action taken by the Receiver which also favors Sender T. Therefore, by the tie breaking rule, once Sender T have been satisfied, Sender $T-1$ has the highest priority.

$$V_{T-1}^{T-1}(\mu) = E_\mu[v^{T-1}(\rho^S(\mu),\mu)] \geqslant E_\mu[v^{T-1}(a',\mu)] = \lim_{k\to\infty} E_{\mu^{n_k}}[v^{T-1}(\rho^S(\mu^{n_k}),\mu^{n_k})]$$
$$= \lim_{k\to\infty} V_{T-1}^{T-1}(\rho^S(\mu^{n_k}),\mu^{n_k}) = \overline{\lim}_{n\to\infty} V_{T-1}^{T-1}(\mu^n)$$

Contradiction. □

Lemma 5

Proof. It is immediate that $cl(\tilde{V}_{T-1}^{T-1}) \leq cl(V_{T-1}^{T-1})$. The remaining part is to show $cl(\tilde{V}_{T-1}^{T-1}) \geq cl(V_{T-1}^{T-1})$. If there is μ^*, $cl(\tilde{V}_{T-1}^{T-1})(\mu^*) < cl(V_{T-1}^{T-1})(\mu^*)$, $\exists\langle\theta_1,\ldots,\theta_m\rangle_{\mu^*}$ s.t. $cl(\tilde{V}_{T-1}^{T-1})(\mu^*) < \sum_{j=1}^m p(\theta_j)V_{T-1}^{T-1}(\theta_j)$. WLOG, suppose $\theta_1 \in \mathcal{G}^T$, and $\{\theta_i\}_{i\neq 1} \subset (\mathcal{G}^T)^c$. Then, $V_{T-1}^{T-1}(\theta_1) = \sum_{j=1}^m p(\eta_j)V_T^{T-1}(\eta_j)$ s.t. $\sigma_T^S(\theta_1) := \langle\eta_1,\ldots,\eta_m\rangle_{\theta_1}$. By Lemma 1, $\eta_j \in (\mathcal{G}^T)^c$, for $\forall j$. Therefore, $\tilde{V}_{T-1}^{T-1}(\eta_j) = V_{T-1}^{T-1}(\eta_j)$, for $\forall j$.

$$cl(\tilde{V}_{T-1}^{T-1})(\mu^*) < p(\theta_1)\sum_{j=1}^m p(\eta_j)\tilde{V}_{T-1}^{T-1}(\eta_j) + \sum_{j=2}^m p(\theta_j)\tilde{V}_{T-1}^{T-1}(\theta_j)$$

Contradiction with the fact that $\{\eta_1,\ldots,\eta_m,\theta_2,\ldots,\theta_m\} \subset (\mathcal{G}^T)^c$. □

Lemma 6

Proof. As $\mathcal{G}^{T-1}$ is the *active* set for $\tilde{V}_{T-1}^{T-1}$, trivially, $\mathcal{G}^{T-1} \supset \mathcal{G}^T$ and on the opposite $(\mathcal{G}^{T-1})^c \subset (\mathcal{G}^T)^c$. It is equal to prove that $(\mathcal{G}^{T-1})^c$ closed, that is, for any $\{\mu^n\} \subset (\mathcal{G}^{T-1})^c$ and $\{\mu^n\} \to \mu$, it can be shown that $\mu \in (\mathcal{G}^{T-1})^c$, i.e. $V_{T-1}^{T-1}(\mu) = cl(V_{T-1}^{T-1})$. Because $cl(V_{T-1}^{T-1})$ is LSC, suppose $\lim_{n\to\infty} V_{T-1}^{T-1}(\mu^n) = \underline{\lim}_{n\to\infty} V_{T-1}^{T-1}(\mu^n) = x$, it suffices to prove that $V_{T-1}^{T-1}(\mu) \geq x \geq cl(V_{T-1}^{T-1})(\mu)$.

Because A is compact, there exists a subsequence $\{\mu^{n_k}\}$ such that $\rho^S(\mu^{n_k}) \to a'$. By the Maximum Theorem and the continuity of V_T^T in $(\mathcal{G})^c$, we have $a' \in A^*(\mu)$ and $V_T^T(\mu) = E_\mu[v^T(a',\mu)]$, which means that a' has been an action that favors Sender T. Additionally, Sender $T-1$ enjoys the highest priority among the remaining senders, s.t.

$$V_T^{T-1}(\mu) \geq E_\mu[v^{T-1}(a',\mu)] = \lim_{k\to\infty} E_{\mu^{n_k}}[v^{T-1}(\rho^S(a^{n_k}),\mu^{n_k})] = \lim_{k\to\infty} V_T^{T-1}(\mu^{n_k}) = x.$$

Because $\mu \in (\mathcal{G}^{T-1})^c \subset (\mathcal{G}^T)^c$, $V_T^{T-1}(\mu) = V_{T-1}^{T-1}(\mu)$. Then $V_{T-1}^{T-1}(\mu) \geq x$. □

12.5 Proposition 9

Proof. The proof is organized in an order as: (4) $\Rightarrow$ (5) $\Rightarrow$ (6) $\Rightarrow$ (3) $\Rightarrow$ (1) $\Rightarrow$ (2). And below I abbreviate terms in the condition part as cond. $(\cdot)$ and terms in the conclusion part as res. $(\cdot)$.

Proof of res.(4)

Sender t's optimal signaling rules are derived from $\tilde{V}_t^t$ and $cl(V_t^t)$ according to cond.(1). For any $\mu \in \mathcal{G}^{t+1}$, $\tilde{V}_t^t(\mu) = -\infty$ which can not serve as an induced posterior in equilibrium. For any $\mu \in \mathcal{G}^t - \mathcal{G}^{t+1}$, $cl(V_t^t)(\mu) > V_t^t(\mu)$. Following the same argument in the proof of Lemma 1, it cannot be an induced posterior in equilibrium.

Proof of res.(5)

The proof is a counterpart of the proof of Lemma 4 with little modification. Because of cond. (2), $\tilde{V}_{t-1}^{t-1}$ is USC in $\mathcal{G}^t$ by definition. For $(\mathcal{G}^t)^c$, the proof follows the same logic with that of Lemma 4 with an additional condition that based on cond. (3), the converging action a' in the proof of Lemma 4 is now one of the Receiver's optimal actions favoring Sender k, $k \geq t+1$. Then $\tilde{V}_{t-1}^{t-1}$ is USC by the same reason why $\tilde{V}_{T-1}^{T-1}$ is USC.

Proof of res.(6)

The proof is a counterpart of the proof of Lemma 5. The key is that for any $\mu \in \Delta(\Omega)$, $cl(V_{t-1}^{t-1})(\mu)$ can be achieved by signaling rules with posteriors in $(\mathcal{G}^t)^c$. Here according to cond.(1) and (4), Sender $t-1$'s value functions have only been transformed within the area $\mathcal{G}^t$ with posteriors outside, such that for any $\theta \in \mathcal{G}^t$, $V_{t-1}^{t-1}(\theta) = \sum_{j=1}^m q_j(\phi_j) V_t^{t-1}(\phi_j) = \sum_{j=1}^m q_j(\phi_j) V_{t-1}^{t-1}(\phi_j)$, where $\phi_j \notin \mathcal{G}^t$ for $\forall j$.

Proof of res.(3)

It is immediate from the definition of $\tilde{V}_{t-1}^{t-1}$.

Proof of res.(1)

The proof is a counterpart of the proof of Lemma 2 where I only need to replace V_T^T and $\mathcal{G}^T$ with V_{t-1}^{t-1} and $\mathcal{G}^{t-1}$, respectively. Furthermore, the converging a' is now one of the Reveiver's optimal actions that also favors Sender k, $k \geq t$, because of cond.(3) and res.(2). Then the similar proof applies.

Proof of res.(2)

The proof is a counterpart of the proof of Lemma 3. Because of cond.(3),

(4) and res.(3), it is obvious that V_T^k is continuous in $(\mathcal{G}^{t-1})^c$, for $\forall k \geq t$.

As for the continuity of V_T^{t-1}, I only need to adopt the proof of Lemma 3 with $\mathcal{G}^T$ and V_T^T replaced by $\mathcal{G}^{t-1}$ and V_T^{t-1}, respectively. Same with the proof of res.(1), the converging a' becomes one of Reveiver's optimal actions that also favors Sender k, $k \geq t$ because of cond.(3) and res.(2). Then the proof of Lemma 3 applies.

□

12.6 Theorem 5

$(\pi' > \pi \Rightarrow \pi' \succ \pi)$

Because $\pi' > \pi$, we can write down π's posterior as convex combinations

$$\lambda_i = \sum_{j=1}^{m_2} r_{ij}\mu_j, \text{ s.t. } r_{ij} \in [0,1], \sum_{j=1}^{m_2} r_{ij} = 1, \forall i, \forall j \tag{19}$$

As π' is *singular*, the sequence of numbers $\{r_{ij}\}_{\forall i, \forall j}$ is unique. Also, by the property of Bayesian plausibility, π and π' share the same average of posteriors,

$$p_1\lambda_1 + \cdots + p_{m_1}\lambda_{m_1} = q_1\mu_1 + \cdots + q_{m_2}\mu_{m_2} \tag{20}$$

Plug Eq. (19) into Eq. (20), obtaining that $\sum_i \sum_j p_i r_{ij}\mu_j = \sum_j q_j\mu_j$. Interchanging the summation signs on the left side of the equation,

$$\sum_{j=1}^{m_2}(\sum_{i=1}^{m_1} p_i r_{ij})\mu_j = \sum_{j=1}^{m_2} q_j\mu_j$$

Because of *singularity* and the uniqueness of those coefficients, for $\forall j = 1, \ldots, m_2$,

$$\sum_{i=1}^{m_1} p_i r_{ij} = q_j \tag{21}$$

To prove *sufficiency* I turn to construct the matrix C such that for $\forall i, j$,

$$a_{ij} = \sum_{k=1}^{m_2} c_{ik} b_{kj} \tag{22}$$

Based on a formula in Kamenica and Gentzkow (2011), one can refer posteriors back to their signal generating mechanisms, such that for $\forall i, j$,

$$a_{ij} = \frac{\lambda_i(\omega_j)p_i}{\mu_0(\omega_j)} \tag{23}$$

$$b_{kj} = \frac{\mu_k(\omega_j)q_k}{\mu_0(\omega_j)} \tag{24}$$

Substitute (23) (24) into (22) and obtain a necessary condition for Eq. (22),

$$\lambda_i(\omega_j)p_i = \sum_{k=1}^{m_2} c_{ik}\mu_k(\omega_j)q_k \tag{25}$$

It could be easily verified that Eq. (25) is satisfied when for $\forall i, j$,

$$c_{ij} = \frac{p_i r_{ij}}{q_j} \tag{26}$$

Specifically, $\sum_i c_{ij} = 1$ according to Eq. (21). Therefore, there exists a matrix C that satisfies the requirement.

$(\pi' \supset \pi \Rightarrow \pi' > \pi)$

Prove by contradiction. For any utility function u and action set A, V_u is a convex function as the maximization over a set of linear functions, which is almost the only restriction on V_u for different u and A. Then, if there exists a $\lambda_k, k \in \{1, \ldots, m_1\}$ lying out of $conv\{\mu_1, \ldots, \mu_{m_2}\}$, it is feasible to find a u and A such that V_u takes on an extremely high value at λ_k but a low value within $conv\{\mu_1, \ldots, \mu_{m_2}\}$. As a result, $\bar{V}_u(\pi') < \bar{V}_u(\pi)$. Because $\bar{V}_u$ is the maximal expected payoff of the entries of payoff vectors, $\bar{V}_u(\pi') < \bar{V}_u(\pi)$ means that it is impossible for π' to be more informative than π.

By Blackwell's theorem, it suffices to conclude Theorem 5 with the two claims: $\pi' > \pi \Rightarrow \pi' \succ \pi$ and $\pi' \supset \pi \Rightarrow \pi' > \pi$.