Journal of Mathematical Economics 107 (2023) 102863

Contents lists available at ScienceDirect

Journal of Mathematical Economics

journal homepage: www.elsevier.com/locate/jmateco

Bayesian persuasion: Reduced form approach

Akhil Vohra [a], Juuso Toikka [b], Rakesh Vohra [c,*]

[a] Department of Economics, Terry College of Business, University of Georgia, United States of America
[b] Department of Business Economics and Public Policy, The Wharton School, University of Pennsylvania, United States of America
[c] Department of Economics & Department of Electrical and Systems Engineering, University of Pennsylvania, United States of America

ARTICLE INFO

Article history:
Received 16 July 2022
Received in revised form 6 February 2023
Accepted 20 May 2023
Available online 1 June 2023
Manuscript handled by Editor C.E. Fabio Maccheronni

Keywords:
Bayesian persuasion
Information design
Mechanism design
Duality

ABSTRACT

We introduce reduced form representations of Bayesian persuasion problems where the variables are the probabilities that the receiver takes each of her actions. These are simpler objects than, say, the joint distribution over states and actions in the obedience formulation of the persuasion problem. This can make a difference in computational and analytical tractability, which we illustrate with two applications. The first shows that with quadratic receiver payoffs, the worst-case complexity scales with the number of actions and not the number of states. If $|\mathcal{A}|$ and $|\mathcal{S}|$ denote the number of actions and states respectively, the worst case complexity of the obedience formulation is $O(|\mathcal{A}||\mathcal{S}|(|\mathcal{S}| + |\mathcal{A}|)^{1.5}L)$ where L is its input size. The worst-case complexity of the reduced form representation is $O(|\mathcal{A}|^{2.5}L)$. In the second application, the reduced form leads to a simple greedy algorithm to determine the maximum value a sender can achieve in any cheap talk equilibrium.

1. Introduction

The model of Bayesian persuasion in Kamenica and Gentzkow (2011) is now the main framework for investigating how a principal can use information rather than carrots and sticks to influence the behavior of an agent.[1] There is an underlying state initially unknown to both principal (called the sender) and agent (called the receiver). The receiver wishes to choose an action whose payoff depends on the unknown state. That action affects the sender's payoffs as well. The state, when realized, is revealed only to the sender. However, *before* the state is realized, the sender *commits* to how much information about the state she will reveal to the receiver. Any information revealed by the sender affects the posterior beliefs of the receiver, thereby affecting the receiver's action choice.[2] Should the sender obfuscate the actual state, and if so, how?

The sender's problem of choosing what information to reveal about the state to maximize her payoff can be formulated in three ways. The first is in terms of choosing a decomposition of the prior distribution over states into a convex combination of possible posterior distributions. This decomposition yields the information structure, that is, the mapping from state to signals that the sender should employ to maximize her expected payoff (e.g., Kamenica and Gentzkow, 2011; Dworczak and Martini, 2019; Doval and Skreta, 2021). The second, called concavification, does not explicitly identify the optimal signal structure. Instead, it characterizes the sender's optimal expected payoff in terms of a concave envelope. For examples, see Lipnowski and Mathevet (2018) and Lipnowski and Ravid (2020). The third assumes that the sender recommends an action as a function of the underlying state. These recommendations must be in the receiver's interest to follow. For this reason we call it the obedience formulation (e.g. Kolotilin, 2018; Dughmi and Xu, 2016; Dughmi et al., 2019; Salamanca, 2021; Galperti and Perego, 2018).

Our paper proposes a reduced-form representation of the obedience formulation. Reduced form representations of optimization problems have proved useful in other settings. See, for example, Che et al. (2013), Epitropou and Vohra (2019), Pai and Vohra (2014), Bertsimas and Niño-Mora (1996) and Queyranne and Schulz (1994). In our case, the reduced form variables are the probabilities with which the receiver takes each of her actions. This bypasses the complications associated with decompositions of distributions or concavification of functions.[3]

We demonstrate the usefulness of the approach with two applications. In each, the sender's preferences are *state independent*. In the first, the receiver cares about matching the state as measured by a quadratic loss function. Suppose $|\mathcal{A}|$ is the number

* Corresponding author.
E-mail addresses: akhil.vohra@uga.edu (A. Vohra), toikka@wharton.upenn.edu (J. Toikka), rvohra@seas.upenn.edu (R. Vohra).

1 See Kamenica (2019) for a survey.

2 A standard alternative interpretation is that the sender does not observe the state either but can design an arbitrary experiment whose result is observed by the receiver, who then takes an action.

3 Mathematically, these are equivalent, but a reduced-form representation may reveal structure obscured by other representations.

https://doi.org/10.1016/j.jmateco.2023.102863
0304-4068/

of actions, $|\mathcal{S}|$ the number of states and L the input size.[4] Then, the number of elementary operations needed to find an optimal solution grows, at worst, as $|\mathcal{A}|^{2.5}L$ (e.g. the worst-case complexity is no more than $O(|\mathcal{A}|^{2.5}L)$). Thus, the complexity does not scale with the number of states, which is an advantage in settings where the number of states far exceeds the number of actions. By comparison, the worst-case complexity of solving the obedience formulation as a linear program is $O(|\mathcal{A}||\mathcal{S}|(|\mathcal{S}|+|\mathcal{A}|)^{1.5}L)$. Note, the input size of both programs is the same.

In the second application, the receiver also benefits from matching the state, but unlike the quadratic loss case, she incurs a fixed, state-dependent cost when she mismatches. While the reduced form representation does not suggest a simple algorithm for solving the persuasion problem, it does yield a simple greedy algorithm to determine the maximum value a sender can achieve in any cheap talk equilibrium.

Section 2 of this paper describes the obedience formulation of the persuasion problem. Section 3 describes the reduced form representation for the first application. Section 4 discusses the cheap talk application.

2. The persuasion problem

We formulate the optimal persuasion problem with a finite number of states and actions as a linear program. Let $\mathcal{S}$ be a finite set of states and $\mathcal{A}$ a finite set of actions. Elements of each are denoted by ω_j and a_i, respectively.

We restrict attention to the so-called *pure persuasion environment* where the sender's (she/her) payoff is independent of the state and depends only on the receiver's (he/him) action.[5] If the receiver chooses action a_i in state ω_j, his payoff is denoted $V_R(\omega_j, a_i)$, and the value to the sender is denoted $V_S(a_i)$. The sender and the receiver share a common prior p over $\mathcal{S}$.

Let $x(\omega_j, a_i)$ be the (joint) probability of the realized state being ω_j and the sender recommending action a_i to the receiver. The sender's optimization problem (see Myerson, 1991 and Bergemann and Morris, 2016) is

$$\max_{x(\omega,a)} \sum_{i=1}^{|\mathcal{A}|}\sum_{j=1}^{|\mathcal{S}|} V_S(a_i)x(\omega_j, a_i)$$

$$\text{s.t.} \sum_{j=1}^{|\mathcal{S}|} V_R(\omega_j, a_i)x(\omega_j, a_i) \geq \sum_{j=1}^{|\mathcal{S}|} V_R(\omega_j, a_k)x(\omega_j, a_i) \text{ for all } a_i \text{ and } a_k \tag{1}$$

$$\sum_{i=1}^{|\mathcal{A}|} x(\omega_j, a_i) = p(\omega_j) \text{ for all } \omega_j \in \mathcal{S} \tag{2}$$

$$x(\omega_j, a_i) \geq 0 \text{ for all } \omega_j \in \mathcal{S} \text{ and } a_i \in \mathcal{A}. \tag{3}$$

Constraints (1) are the obedience constraints (hereafter referred to as OC) that ensure that it is in the receiver's interest to follow the sender's recommendation.

Constraints (2) ensure that the total probability weight assigned to actions recommended in state ω_j matches the prior probability of state ω_j being realized.

Dughmi et al. (2019), Salamanca (2021) and Galperti and Perego (2018) use duality and complementary slackness to characterize the optimal solution of (1)–(3). Our point is that *other* formulations of the persuasion problem can sometimes be more useful.

[4] Given a linear program with constraint matrix $M = \{m_{ij}\}$, objective function vector c and right hand side vector b, the input size is $\sum_{ij}\log|m_{ij}| + \sum_j \log|c_j| + \sum_i \log|b_i|$.

[5] This is an oft-studied case in the literature, see for example Brocas and Carrillo (2007) and Lipnowski et al. (2020).

3. Pure persuasion: Receiver cares about posterior mean

Assume that $\mathcal{S}$ is a finite set of distinct real numbers with at least two elements and that the optimal action for the receiver is not the same in all states. Suppose also that the receiver's preferences over actions depend only on the posterior mean of the state. First, we characterize of the class of receiver payoffs that satisfy this condition. Then, we offer a reduced form representation of the associated persuasion problem.

Given a posterior $p \in \Delta(\mathcal{S})$, the receiver's preferences over mixed actions are given by the expected utility $\sum_{a\in\mathcal{A}}\sigma(a)u(p, a)$, where $u(p, a) = \sum_{\omega\in\mathcal{S}} p(\omega)V_R(\omega, a)$. We say that *the receiver's preferences depend only on the posterior mean* if for all posteriors $p, q \in \Delta(\mathcal{S})$ such that $\sum_\omega p(\omega)\omega = \sum_\omega q(\omega)\omega$, there exist constants $\alpha > 0$ and β such that $u(q, a) = \alpha u(p, a) + \beta$ for all a. That is, the receiver's preferences over mixed actions are equivalent in the usual sense given p or q.

The following is, we think, folklore, but as we were unable to find a reference, we include it for completeness.

Theorem 3.1. *The receiver's preferences depend only on the posterior mean if and only if there exist functions $f, g : \mathcal{A} \to \mathbb{R}$ and $h : \mathcal{S} \to \mathbb{R}$ such that:*[6]

$$V_R(\omega, a) = f(a) + g(a)\omega + h(\omega). \tag{4}$$

The proof of Theorem 3.1 appears in the appendix. Armed with it, we may assume $V_R(\omega_j, a_i) = f(a_i) + g(a_i)\omega_j$. The obedience constraint needed to enforce action a_i is:

$$\begin{aligned}
&\sum_{\omega_j\in\mathcal{S}} V_R(\omega_j, a_i)x(\omega_j, a_i) \geq \sum_{\omega_j\in\mathcal{S}} V_R(\omega_j, a_{i'})x(\omega_j, a_i)\\
&\Rightarrow \sum_{\omega_j\in\mathcal{S}}[f(a_i) + g(a_i)\omega_j]x(\omega_j, a_i) \geq \sum_{\omega_j\in\mathcal{S}}[f(a_{i'}) + g(a_{i'})\omega_j]x(\omega_j, a_i)\\
&\Rightarrow [g(a_i) - g(a_{i'})]\sum_{\omega_j\in\mathcal{S}}\omega_j x(\omega_j, a_i) \geq [f(a_{i'}) - f(a_i)]\sum_{\omega_j\in\mathcal{S}} x(\omega_j, a_i).
\end{aligned} \tag{5}$$

Depending on the sign of $\frac{f(a_{i'})-f(a_i)}{g(a_i)-g(a_{i'})}$, Eq. (5) yields either an upper or lower bound on $\frac{\sum_{\omega_j\in\mathcal{S}}\omega_j x(\omega_j,a_i)}{\sum_{\omega_j\in\mathcal{S}} x(\omega_j,a_i)}$. For each a_i let B_i be the set of actions $a_{i'}$ such that:

$$\frac{\sum_{\omega_j\in\mathcal{S}}\omega_j x(\omega_j, a_i)}{\sum_{\omega_j\in\mathcal{S}} x(\omega_j, a_i)} \geq \frac{f(a_{i'}) - f(a_i)}{g(a_i) - g(a_{i'})}.$$

Similarly, let U_i be the set of actions $a_{i'}$ such that:

$$\frac{\sum_{\omega_j\in\mathcal{S}}\omega_j x(\omega_j, a_i)}{\sum_{\omega_j\in\mathcal{S}} x(\omega_j, a_i)} \leq \frac{f(a_{i'}) - f(a_i)}{g(a_i) - g(a_{i'})}$$

Let $a_{i_B} \in B_i$ be the index that maximizes $\frac{f(a_{i'})-f(a_i)}{g(a_i)-g(a_{i'})}$. Similarly, let $a_{i_U} \in U_i$ be the index that minimizes $\frac{f(a_{i'})-f(a_i)}{g(a_i)-g(a_{i'})}$. Hence, Eq. (5) yields:

$$\frac{f(a_{i_U}) - f(a_i)}{g(a_i) - g(a_{i_U})} \geq \frac{\sum_{\omega_j\in\mathcal{S}}\omega_j x(a_i, \omega_j)}{\sum_{\omega_j\in\mathcal{S}} x(a_i, \omega_j)} \geq \frac{f(a_{i_B}) - f(a_i)}{g(a_i) - g(a_{i_B})}.$$

Thus, the persuasion problem reduces to:

$$\max \sum_{a_i}\sum_{\omega_j\in\mathcal{S}} x(\omega_j, a_i)V_S(a_i) \tag{6}$$

[6] Of course, the term $h(\omega)$ does not affect the receiver's behavior, so for any V_R satisfying (4), there exists a behaviorally equivalent payoff function with $h \equiv 0$.

$$\text{s.t. } \frac{f(a_{i_U})-f(a_i)}{g(a_i)-g(a_{i_U})} \geq \frac{\sum_{\omega_j\in\mathcal{S}} \omega_j x(\omega_j,a_i)}{\sum_{\omega_j\in\mathcal{S}} x(\omega_j,a_i)} \geq \frac{f(a_{i_B})-f(a_i)}{g(a_i)-g(a_{i_B})} \; \forall a_i \tag{7}$$

The reduced form representation is formulated in terms of the marginal distribution over the receiver's action rather than its joint distribution with the state of the world. In other words, we would like to reformulate problem (6)–(7) in terms of the variables $z_i = \sum_{j=1}^{|\mathcal{S}|} x(\omega_j, a_i)$. To do this, we project the polyhedron (7) onto the "z-space". The goal is to provide a succinct characterization of this projection.

To illustrate, we use a standard payoff function where the receiver's payoff depends on how close the action is to the state as measured by quadratic loss. The setting is canonical, see for example Dworczak and Martini (2019) and Kolotilin (2018). Hence, $V_R(\omega_j, a_i) = -(a_i - \omega_j)^2$. Such a payoff function is a special case of (4). This can be seen by taking $f(a) = -\frac{1}{2}a^2$ and $g(a) = a$. Then, $V_R(\omega, a) = a\omega - \frac{1}{2}a^2$, which can be written equivalently as $-\frac{1}{2}(a-\omega)^2 + \frac{1}{2}\omega^2$. Of course, the term $\frac{1}{2}\omega^2$ does not affect preferences over $\mathcal{A}$, and hence it can be omitted—as is often done—if we are only interested in the optimal choice of a.

Without loss, we order the states and actions in $\mathcal{S}$ and $\mathcal{A}$ in increasing order:

$$\omega_1 < \omega_2 < \ldots < \omega_{|\mathcal{S}|}$$

$$a_1 < a_2 < \ldots < a_{|\mathcal{A}|}$$

The persuasion problem in its obedience formulation simplifies, via (7) to the following:

$$\max \sum_{i=1}^{|\mathcal{A}|}\sum_{j=1}^{|\mathcal{S}|} V_S(a_i)x(\omega_j, a_i) \tag{8}$$

$$\frac{a_i + a_{i+1}}{2} \geq \frac{\sum_{j=1}^{|\mathcal{S}|}\omega_j x(\omega_j,a_i)}{\sum_{j=1}^{|\mathcal{S}|} x(\omega_j,a_i)} \geq \frac{a_i+a_{i-1}}{2} \;\; \forall i \in \{2,\ldots,|\mathcal{A}|-1\} \tag{9}$$

$$\frac{a_1+a_2}{2} \geq \frac{\sum_{j=1}^{|\mathcal{S}|}\omega_j x(\omega_j,a_i)}{\sum_{j=1}^{|\mathcal{S}|} x(\omega_j,a_i)} \tag{10}$$

$$\frac{\sum_{j=1}^{|\mathcal{S}|}\omega_j x(\omega_j,a_{|\mathcal{A}|})}{\sum_{j=1}^{|\mathcal{S}|} x(\omega_j,a_{|\mathcal{A}|})} \geq \frac{a_{|\mathcal{A}|}+a_{|\mathcal{A}|-1}}{2} \tag{11}$$

$$\sum_{i=1}^{|\mathcal{A}|} x(\omega_j, a_i) = p(\omega_j) \;\; \forall \omega_j \in \mathcal{S} \tag{12}$$

$$x(\omega_j, a_i) \geq 0 \;\; \forall a_i \in \mathcal{A}, \omega_j \in \mathcal{S}. \tag{13}$$

This formulation has $|\mathcal{A}||\mathcal{S}|$ variables and $|\mathcal{A}|+|\mathcal{S}|$ constraints.

In Dworczak and Martini (2019) and Kolotilin (2018), the set of states and actions are intervals with $\mathcal{S} \subset \mathcal{A}$. The problem is formulated so that the variable is a distribution over the posterior expected state. When the receiver's preferences satisfy quadratic loss, states can be relabeled to equal the receiver's optimal actions. Thus, the relevant variable becomes the distribution over the receiver's actions. When $\mathcal{S} \subset \mathcal{A}$ this relabeling step is straightforward, but we do not impose this assumption. Further, this observation only shows that a formulation in terms of the distribution over the receiver's actions is *equivalent* rather than *better* in a precise sense. As noted in the introduction, the reduced form representation has a lower worst-case complexity than the obedience formulation.

Theorem 3.2. *For each $\omega_r \in \mathcal{S}$ let*

1. $U^+(\omega_r) = \{i : \frac{a_i+a_{i+1}}{2} > \omega_r\}$
2. $B^+(\omega_r) = \{i : \frac{a_i+a_{i+1}}{2} \leq \omega_r\}$
3. $U^-(\omega_r) = \{i : \frac{a_i+a_{i-1}}{2} > \omega_r\}$
4. $B^-(\omega_r) = \{i : \frac{a_i+a_{i-1}}{2} \leq \omega_r\}$

The persuasion problem (8)–(13) can be expressed as

$$\max_{z_1,\ldots,z_{|\mathcal{A}|}} \sum_{i=1}^{|\mathcal{A}|} V_S(a_i)z_i$$

$$\text{s.t. } \omega_r \sum_{i\in B^-(\omega_r)\cup\{1\}} z_i + \sum_{i\in U^-(\omega_r)} \frac{(a_i+a_{i-1})z_i}{2} \leq \sum_j \max\{\omega_r,\omega_j\}p(\omega_j) \;\; \forall 2 \leq r \leq |\mathcal{S}| \tag{14}$$

$$\omega_r \sum_{i\in U^+(\omega_r)\cup\{|\mathcal{A}|\}} z_i + \sum_{i\in B^+(\omega_r)} \frac{(a_i+a_{i+1})}{2} z_i \geq \sum_j \min\{\omega_j,\omega_r\}p(\omega_j) \;\; \forall 1 \leq r \leq |\mathcal{S}|-1 \tag{15}$$

$$\sum_{i\in\mathcal{A}} z_i = 1 \tag{16}$$

$$z_i \geq 0 \;\; \forall i \in \mathcal{A}. \tag{17}$$

The proof of Theorem 3.2 appears in the appendix. Observe, the number of variables in formulation (14)–(17) depends on $|\mathcal{A}|$ only. Ostensibly, the number of constraints depends on $|\mathcal{S}|$ but many of these will be redundant. To see why, suppose

$$\frac{a_i+a_{i-1}}{2} \leq \omega_j < \omega_{j+1} \leq \frac{a_{i+1}+a_i}{2}.$$

Then, $B^-(\omega_j) = B^-(\omega_{j+1})$ and $U^-(\omega_j) = U^-(\omega_{j+1})$.

In many papers it is common to assume that $\mathcal{A} = \mathcal{S}$ and $a_i = \omega_i = i$ for all i. In this case, the problem (8)–(13) can be expressed as:

$$\max_{z_1,\ldots,z_{|\mathcal{A}|}} \sum_{i=1}^{|\mathcal{A}|} V_S(i)z_i$$

$$\text{s.t.} \quad z_1 + \sum_{i\geq 2}(i-0.5)z_i \leq \sum_i ip(i) \tag{18}$$

$$\sum_{i\in\mathcal{A}} \max\{(i-0.5), r\}z_i \leq \sum_{i\in\mathcal{A}} \max\{i,r\}p(i) \;\; \forall r \geq 2 \tag{19}$$

$$\sum_{i\in\mathcal{A}} \min\{(i+0.5), r\}z_i \geq \sum_{i\in\mathcal{A}} \min\{i,r\}p(i) \;\; \forall r \leq |\mathcal{A}|-1 \tag{20}$$

$$\sum_{i\leq|\mathcal{A}|-1} (i+0.5)z_i + |\mathcal{A}|z_{|\mathcal{A}|} \geq \sum_i ip(i) \tag{21}$$

$$z_i \geq 0 \;\; \forall i \in \mathcal{A}. \tag{22}$$

Constraints (18) and (21) are the discrete analogs of the following:

$$\int_0^1 xz(x)dx = \int_0^1 xp(x)dx.$$

In words, the expected action must equal the expected state. Also, note the absence of (16). This is because it is implied by the other constraints. If we choose $r = |\mathcal{A}|$ in (19), this yields $\sum_{i\in\mathcal{A}} z_i \leq 1$. If we choose $r = 1$ in (20) it yields $\sum_{i\in\mathcal{A}} z_i \geq 1$.

To interpret (19), it is helpful to consider its 'continuous' analog. Suppose $\mathcal{A} = \mathcal{S} = [0, 1]$. Then, (19) can be rendered as:

$$\int_0^1 \max\{x, r\}z(x)dx \le \int_0^1 \max\{x, r\}p(x)dx \ \forall r \in [0, 1].$$

Therefore, the random variable associated with the density $z(x)$ is below the random variable associated with the density $p(x)$ in the increasing convex order (see chapter 3 of Shaked and Shanthikumar, 2007). One consequence is that the variance of the distribution over actions is smaller than the variance of the distribution over the states. In effect, the sender is 'rewarding' the receiver with lower variance in return for taking an action that is more preferred by the sender. Quadratic loss preferences render the receiver risk-averse. Thus, they are willing to trade off a higher mean for lower variance.

As the receiver chooses the action closest to the posterior mean, when the set of actions exactly matches the set of states, a feasible z is equivalent to a distribution over posterior means. The problem of characterizing distributions over posterior means was explored by Gentzkow and Kamenica (2016). They characterize the distributions via its integral. Since the integral of a cumulative distribution function is convex, the problem of selecting a distribution over actions is the same as selecting a distribution over posterior means, which in turn is the same as selecting a convex function from amongst an 'interval' of convex functions.[7] However, there is no characterization of these convex functions. Furthermore, no explicit formulation is given of this optimization problem, as the decision variables are not the z's themselves but a linear functional of them. Thus, our Theorem 3.2 is of particular note in that it provides a simple, explicit characterization of the feasible z's for *any* set of actions and states. Example 1 below illustrates an advantage of our formulation compared to Gentzkow and Kamenica (2016).

As observed above, the size of the reduced form representation scales with the number of actions, $|\mathcal{A}|$ only. In contrast, the size of the obedience formulation scales with $|\mathcal{A}||\mathcal{S}|$. This has implications for the worst-case complexity of solving the persuasion problem.

Theorem 3.3. *Suppose L is the input size of formulation (14)–(17). The worst-case complexity of solving formulation (14)–(17) is no more than $O(|\mathcal{A}|^{2.5}L)$.*

Proof. Vaidya (1989) provides a deterministic algorithm that will solve a linear program with n variables, m constraints and input size L in time $O(n(n+m)^{1.5}L)$. This is not the fastest known algorithm; see Cohen et al. (2019) for example. However, it does admit a succinct complexity bound. Now, formulation (14)–(17) has $|\mathcal{A}|$ variables and, as argued earlier, the same number of constraints. ■

By comparison, the obedience formulation has $|\mathcal{A}||\mathcal{S}|$ variables and $O(|\mathcal{A}|+|\mathcal{S}|)$ constraints. Therefore, the worst-case complexity of solving the obedience formulation as a linear program (which will have the same input size as (8)–(13)) is $O(|\mathcal{A}||\mathcal{S}|(|\mathcal{S}| + |\mathcal{A}|)^{1.5}L)$. Theorem 3.2 indicates the advantage of our approach.

To illustrate the use of reduced form representation, we provide a simple example.

Example 1. We now provide a contrast with the three action example in Gentzkow and Kamenica (2016). They assume the sender prefers increasing actions and are only able to provide an explicit characterization under a uniform prior. We show for *any* sender preferences and prior how elementary manipulations will reduce an instance of (18)–(22) to an optimization problem involving a *single* variable.

Suppose $\mathcal{A} = \mathcal{S} = \{1, 2, 3\}$. Problem (18)–(22) is:

$$\max V_S(1)z_1 + V_S(2)z_2 + V_S(3)z_3$$

$$\text{s.t. } z_1 + 1.5z_2 + 2.5z_3 \le p(1) + 2p(2) + 3p(3)$$

$$2z_1 + 2z_2 + 2.5z_3 \le \sum_{j=1}^{3} \max\{j, 2\}p(j)$$

$$1.5z_1 + 2z_2 + 2z_3 \ge \sum_{j=1}^{2} \min\{j, 2\}p(j)$$

$$1.5z_1 + 2.5z_2 + 3z_3 \ge p(1) + 2p(2) + 3p(3)$$

$$z_1 + z_2 + z_3 = 1$$

$$z_1, z_2, z_3 \ge 0$$

Using the constraint $z_1 + z_2 + z_3 = 1$ we can simplify the constraints to

$$0.5z_2 + 1.5z_3 \le p(2) + 2p(3)$$

$$0.5z_3 \le p(3)$$

$$0.5(z_2 + z_3) \ge 0.5 - p(1)$$

$$z_2 + 1.5z_3 \ge p(2) + 2p(3) - 0.5$$

$$z_1 + z_2 + z_3 = 1$$

$$z_1, z_2, z_3 \ge 0$$

The second of these constraints is redundant.

Eliminating z_2, we obtain:

$$\max V_S(1)z_1 + V_S(2)(1 - z_1 - z_3) + V_S(3)z_3$$

$$\text{s.t. } z_3 \le p(2) + 2p(3) + 0.5z_1 - 0.5$$

$$z_1 \le 2p(1)$$

$$z_3 \ge 2p(2) + 4p(3) - 3$$

$$0 \le z_3 \le 1 - z_1$$

Hence, $z_3 = \min\{1 - z_1, p(2) + 2p(3) + 0.5z_1 - 0.5\}$. So, our problem reduces to the following:

$$V_S(2) + \max[V_S(1) - V_S(2)]z_1 + [V_S(3) - V_S(2)] \min\{1 - z_1, p(2) + 2p(3) + 0.5z_1 - 0.5\}$$

$$\text{s.t. } 1 - 2p(2) - 4p(3) \le z_1 \le 2p(1)$$

$$0 \le z_1 \le 1$$

It is common in the literature to assume that both the sender's and receiver's preferences depend on the posterior mean only (e.g. Dworczak and Martini, 2019). Hence, one may wonder whether our approach would extend to this case. For a specific functional form of the sender's payoffs, yes. The analysis is outlined in the appendix (see Theorem 8.1).

[7] One end of this interval corresponds to the distribution over posterior means induced by the most uninformative signal structure and the other induced by the most informative signal.

4. Cheap talk

Our reduced-form approach to the Persuasion problem has an added benefit in that the formulation allows one to easily compute the value of sender commitment. An important assumption in the persuasion problem is that the sender can commit ex-ante to an action recommendation policy. What if the sender cannot commit? Then, we are in the classical cheap talk environment. In the cheap talk version of the pure persuasion problem, the sender chooses a signal structure (with some fixed, large set of signals), and the receiver chooses her strategy (mapping from signals to actions) *simultaneously*.[8] We use Lipnowski and Ravid (2020) to link the reduced-form formulation of the constraints in the persuasion problem to the set of achievable equilibrium payoffs in a cheap talk game. We show that a reduced form representation of the pure persuasion problem can also be used to characterize the maximum payoff a sender can achieve in the cheap talk version of the problem.

To formalize this connection, we will need the following:

Definition 4.1. *The sender is said to secure a payoff Q under information policy x if $V_S(a) \geq Q$ for every action $a \in \mathcal{A}$ recommended with positive probability under x (i.e., for all $a \in \mathcal{A}$ such that $\sum_{j=1}^{|\mathcal{S}|} x(\omega_j, a) > 0$). A payoff Q can be secured if the sender can secure it under some information policy.*

Theorem 4.2 (*Lipnowski and Ravid, 2020*)**.** *Let $\bar{a}$ be the receiver's best-response action under the prior. If $Q \geq V_S(\bar{a})$ can be secured, then there is an equilibrium in the cheap talk game that yields Q to the sender.*

Lipnowski and Ravid (2020) show that the set of sender equilibrium payoffs in the cheap talk game is equivalent to the set of "securable" payoffs for the sender in the corresponding persuasion setting. Critical for our analysis is the *interpretation* of securability. A securable payoff is the payoff associated with the worst-case realized action in an information policy. To characterize securable payoffs, it suffices to check which actions information policies can induce. Thus, for any subset of actions $A \subset \mathcal{A}$, define the following function f:

$$f(A) = \max\{\sum_{a_i \in A} \sum_{\omega_j \in \mathcal{S}} x(\omega_j, a_i) : \text{ s.t. } (1)\text{–}(3)\}.$$

The optimal value of this program is the maximum frequency with which the sender can have the receiver take actions in A under any information structure satisfying the obedience constraints. Notice, $f(A) = 1$ implies there is some information policy such that the only actions induced by any signal realization are actions in A. In particular, consider the action a^* that is optimal for the receiver under the prior. Then $f(\{a^*\}) = 1$ and $V_S(a^*)$ is securable. Thus, there is an equilibrium in the cheap talk game where the sender receives a payoff of $V_S(a^*)$, namely the babbling one.

Assume now, without loss of generality, that the actions are labeled in the order of decreasing payoff to the sender, i.e., $V_S(a_1) > V_S(a_2) > \ldots > V_S(a_{|\mathcal{A}|})$.

Theorem 4.3. *Let $k^* = \min\{k : f(\{1 \ldots, k\}) = 1\}$. Then, there is a cheap-talk equilibrium that yields payoff $V_S(a_{k^*})$ to the sender. Furthermore, this is the maximum payoff the sender can receive in any cheap talk equilibrium.*

Proof. By definition of k^*, $V_S(a_{k^*})$ can be secured. Now, for any action a_k, $k < k^*$, suppose $V_S(a_k)$ could be secured. Then, there is an information policy such that for any signal, the induced action must be from the set $\{a_1, \ldots, a_k\}$. However, this would mean $f(\{a_1, \ldots, a_k\}) = 1$, which contradicts the definition of k^*.

Finally, let $\bar{a}$ be the receiver's best-response action under the prior, with no additional information. It must be that $V_S(a_{k^*}) \geq V_S(\bar{a})$ since $f(\{\bar{a}\}) = 1$. Applying Theorem 4.2 yields the desired result. ■

Determining the achievable payoff in any cheap talk equilibrium is conceptually straightforward. Starting with action a_1, compute the maximum probability with which one can induce the receiver to play a_1. Continue greedily, adding actions into the set A until the sender can induce the receiver to play only actions in A.

Theorem 4.3 and its proof demonstrate that the securable payoffs can be entirely characterized by the function f. A payoff of Q can be secured if and only if $f(\{a_1, \ldots, a_k\}) = 1$ for some k such that $V(a_k) \geq Q$. The reduced form representation of the constraints of the Bayesian Persuasion problem allows us to view information policies solely through the distribution over actions. Since the actions determine the value to the sender, computing the possible securable payoffs is greatly simplified when the constraints are expressed in such a reduced form.

This procedure does have a close connection to the concavification approach. Let Π_i denote the set of posteriors that induce action a_i as a best response for the receiver. The following is an alternative characterization of k^*.

Theorem 4.4. *Let $k^* = \min\left\{k : p \in conv\left(\cup_{i=1}^{k} \Pi_i\right)\right\}$. There is a cheap-talk equilibrium that yields payoff $V_S(a_{k^*})$ to the sender. Furthermore, this is the maximum payoff the sender can receive in any cheap talk equilibrium.*

Proof. A distribution of posteriors is feasible if and only if there exists a convex combination of said posteriors equal to the prior. By Theorem 4.3, $V_S(a_{k^*})$ is the maximal securable payoff, which means there is an information policy that induces only actions $a \in \{a_1, \ldots, a_{k^*}\}$ to be played. Hence, the prior $p \in conv\left(\cup_{i=1}^{k^*} \Pi_i\right)$.

Now, suppose $p \in conv\left(\cup_{i=1}^{k} \Pi_i\right)$ for $k < k^*$. This means there is an information policy that induces only actions $a \in \{a_1, \ldots, a_k\} \implies f(\{a_1, \ldots, a_k\}) = 1$. This contradicts the definition of k^*. ■

4.1. Application

The receiver enjoys a benefit $b_{\omega_j} > 0$ if she "matches the state" ω_j and bears cost $-c_{\omega_j} < 0$ if she does not. Formally, for each $\omega \in \mathcal{S}$, there exists a unique $a \in \mathcal{A}$ such that $V_R(a, \omega) = b_\omega$. Moreover, if $V_R(a, \omega) = b_\omega$ then $V_R(a, \omega') = -c_{\omega'}$ for all $\omega' \neq \omega$. In words, no action is optimal for the receiver at more than one state. Hence, it suffices to restrict attention to the case where $|\mathcal{A}| = |\mathcal{S}| = n > 0$ as assumed above. The sets $\mathcal{A}$ and $\mathcal{S}$ will be represented by $\{a_1, \ldots, a_n\}$ and $\{\omega_1, \ldots, \omega_n\}$, respectively. Without loss, we assume action a_i is optimal in state ω_i and, as above, $V_S(a_1) > V_S(a_2) > \ldots > V_S(a_n)$. If the receiver selects a_i in state ω_i, we say that she matches the state. The matching utility example of Bergemann et al. (2018) is a special case where $c_{\omega_j} = 0$ for all ω_j. We provide further examples to motivate this specification.

Example 2. An incumbent politician must implement a policy to combat an impending crisis or adapt to a new state of affairs. In other words, the status quo, no longer tenable, will be replaced

[8] In contrast to the previous section, the states and actions need not be real numbers, and the receiver's payoff need not be given by quadratic loss.

by a state $\omega_j \in \mathcal{S}$, and the politician must react to the new environment appropriately.

The politician receives information from an ideological think tank. The think tank will commission studies and research efforts to inform the politician about the state. These studies are represented by a signal $\psi : \mathcal{S} \to \Delta S$, where S is the signal space. Once the politician observes the choice of ψ and the realized signal, she selects a policy $a \in \mathcal{A}$ to implement.

The think tank has preferences over the implemented policies. The politician's payoff depends on whether she matches the state. For each state $\omega_j \in \mathcal{S}$, there is an ideal policy $a_j \in \mathcal{A}$. If she implements a_j in state ω_j, she increases her chance of being re-elected by b_{ω_j}. If she implements policy $a \neq a_j$, and the state turns out to be ω_j, she decreases that chance by c_{ω_j}. Whether or not the politician is re-elected is irrelevant to the think tank.

Remark. Implicit is that none of the policies in $\mathcal{A}$ are "outlandish" in the sense of being much worse than the others.

Example 3. A company must select a technology that will be adopted firm-wide. The set of possible technologies is given by $\mathcal{A}$. Because of technological obsolescence, it is likely that only one of the technologies in $\mathcal{A}$ will become dominant, while the others will become antiquated. That is, if the company adopted technology $a \in \mathcal{A}$ and technology $a' \in \mathcal{A}$ became dominant, it would need to replace all of its current technology, and its employees would need time to learn how to use a'. In other words, the company would incur a switching cost.

A seller has an inventory of each of the technologies in $\mathcal{A}$. It has preferences over the technologies it wants to sell. The seller can commit to a signaling policy (studies, research, polls, surveys, etc.) to inform the company about which technologies will become obsolete. In this setting, the states in $\mathcal{S}$ correspond to the technology that becomes the "winner".

Suppose the company selects action $a_j \in \mathcal{A}$, meaning it adopts technology a_j. If the state turns out to be ω_j, it incurs benefit b_{ω_j}. If the state is $\omega_k \neq \omega_j$, it incurs a switching cost of c_{ω_k}. In other words, state ω_k represents the setting where technology a_k becomes dominant, and so the company must now switch to the dominant technology.

To solve for the equilibrium payoff in a cheap talk equilibrium, we start by describing the constraints that characterize *feasible information policies:*

$$-\sum_{j\neq i} c_{\omega_j} x(a_i,\omega_j) + b_{\omega_i} x(a_i,\omega_i) \geq -\sum_{j\neq k} c_{\omega_j} x(a_i,\omega_j) + b_{\omega_k} x(a_i,\omega_k)\ \forall i,k \in \{1,\dots,n\} \tag{23}$$

$$\sum_{i=1}^{n} x(a_i,\omega_j) = p(\omega_j) \text{ for all } j \in \{1,\dots,n\} \tag{24}$$

$$x(a_i,\omega_j) \geq 0 \text{ for all } i,j \in \{1,\dots,n\}. \tag{25}$$

Constraints (23) are the obedience constraints. We focus on the maximum frequency with which the sender can have the receiver take actions in A under any information structure satisfying the obedience constraints.

Theorem 4.5. *For any subset of actions $A \subset \mathcal{A}$ let*

$$f(A) = \max\{\sum_{a_i\in A}\sum_{\omega_j\in\mathcal{S}} x(\omega_j,a_i) : \text{ s.t. } ((23)\text{–}(25))\}.$$

Then, $f(A) = \sum_{a_j\notin A} \min\{p(\omega_j), \sum_{a_i\in A} \frac{b_{\omega_i}+c_{\omega_i}}{b_{\omega_j}+c_{\omega_j}} p(\omega_i)\} + \sum_{a_i\in A} p(\omega_i)$.

Proof. The OC simplifies to:

$$(b_{\omega_i}+c_{\omega_i})x(\omega_i,a_i) - (b_{\omega_k}+c_{\omega_k})x(\omega_k,a_i) \geq 0 \text{ for all } i,k \in \{1,\dots,n\}.$$

For notational convenience, set $\alpha_i = b_{\omega_i} + c_{\omega_i}$ for all $i \in \{1,\dots,n\}$. Hence, the constraints of the persuasion problem can be expressed as follows:

$$-\alpha_i x(\omega_i,a_i) + \alpha_k x(\omega_k,a_i) \leq 0 \text{ for all } i,k \in \{1,\dots,n\} \tag{26}$$

$$\sum_{i=1}^{|\mathcal{A}|} x(\omega_j,a_i) = p(\omega_j) \text{ for all } j \in \{1,\dots,n\} \tag{27}$$

$$z_{a_i} - \sum_{j=1}^{|\mathcal{S}|} x(\omega_j,a_i) = 0 \text{ for all } i \in \{1,\dots,n\} \tag{28}$$

$$x(\omega_j,a_i) \geq 0 \text{ for all } i,j \in \{1,\dots,n\}. \tag{29}$$

We wish to eliminate the x variables. To do so, we interpret the system (26)–(29) in terms of a network flow problem. Each $\omega_j \in \mathcal{S}$ corresponds to a supply node with supply $p(\omega_j)$. Each $a_i \in \mathcal{A}$ corresponds to a demand node with demand z_{a_i}. Any supply node can serve any demand node. However, there is a side constraint:

$$x(\omega_k,a_i) \leq \alpha_i\alpha_k^{-1} x(\omega_i,a_i).$$

For each i, fix the value of $x(\omega_i,a_i)$ at some $\Delta_i \leq p(\omega_i)$. Then, the constraints for a feasible flow must satisfy:

$$x(\omega_j,a_i) \leq \alpha_i\alpha_j^{-1}\Delta_i \text{ for all } i,j \in \{1,\dots,n\}, j \neq i$$

$$\sum_{a_i\neq a_j} x(\omega_j,a_i) = p(\omega_j) - \Delta_j \text{ for all } j \in \{1,\dots,n\}$$

$$z_{a_i} - \Delta_i - \sum_{\omega_j\neq\omega_i} x(\omega_j,a_i) = 0 \text{ for all } i \in \{1,\dots,n\}$$

$$x(\omega_j,a_i) \geq 0 \text{ for all } i,j \in \{1,\dots,n\}.$$

Now, these equations describe a standard flow problem with capacity constraints on the arc flows. Each supply node has supply $p(\omega_j) - \Delta_j$, and each demand node demands $z_{a_i} - \Delta_i$. By Gale's demand theorem (see Gale, 1957), this flow problem is feasible if and only if for all $A \subseteq \mathcal{A}$, we have:

$$\sum_{a_i\in A} z_{a_i} \leq \sum_{j\notin A} \min\{p(\omega_j) - \Delta_j, \sum_{i\in A} \alpha_i\alpha_j^{-1}\Delta_i\} + \sum_{i\in A} p(\omega_i). \tag{30}$$

In words, the total demand in any subset A of demand nodes cannot exceed the total supply of all supply nodes that service them.

Observe that the right hand side of (30) is maximized when we set $\Delta_j = 0$ for all $a_j \notin A$ and $\Delta_i = p(\omega_i)$ for all $a_i \in A$. ■

Rather than focusing on the specific values of $x(a_i,\omega_j)$ or the signal structure, the sender's problem reduces to one of "how much flow can she transmit to action a_i?"

For each index i let $A^i = \{a_1,\dots,a_i\}$ and set $k^* = \min\left\{i : f(A^i) = 1\right\}$. From the definition of $f(A)$, it follows that k^* is the smallest index such that:

$$\max_{j\geq k^*+1} (b_{\omega_j} + c_{\omega_j})p(\omega_j) \leq \sum_{i=1}^{k^*} (b_{\omega_i} + c_{\omega_i})p(\omega_i) \leq \min_{j\leq k^*} (b_{\omega_j} + c_{\omega_j})p(\omega_j).$$

By Theorem 4.3, the maximum achievable payoff in a cheap talk equilibrium is $V_S(a_{k^*})$.

Example 4. Returning to the matching utility example of Bergemann et al. (2018), our analysis above provides an easy way of determining the maximal sender payoff. The network flow interpretation of f allows for a simple characterization of the signal structure. Using the language of the proof of Theorem 4.5, as each action is associated with a state, sending flow from supply node ω_j to demand node a_i represents the partial pooling of state ω_j with ω_i. Therefore, our computation of the closed-form expression of f in Theorem 4.5 determines which states should be pooled.

Let $A = \{a_1, \ldots, a_{k^*}\}$, and consider a signal space with k^* elements, $\{s_1, \ldots, s_{k^*}\}$. The sender will take the states ω_j for $j > k^*$ and pool them with the states ω_i for $i \leq k^*$. This creates k^* pools, one for each ω_i for $i \leq k^*$.

Conceptually, the sender will let the receiver know which pool the state is in: for each $i \leq k^*$, the sender will send a signal s_i for all states that are pooled with ω_i. The conditional state probabilities are given by the fraction of flow into the demand node a_i.

Formally, the signal structure π is defined, as follows:

1. For each $i > k^*$, define $\hat{p}^i(\omega_j) = \max\left\{p(\omega_j) - \sum_{k=1}^{i-1} p(\omega_k), 0\right\}$
2. Set $\pi(s_i) = p(\omega_i) + \sum_{a_j \notin A} \min\left\{p(\omega_i), \hat{p}^i(\omega_j)\right\}$
3. For each $j > k^*$, set $\pi(\omega_j|s_i) = \dfrac{\mathbb{I}_{\hat{p}^i(\omega_j) \geq p(\omega_i)}}{1 + \sum_{a_k \notin A} \mathbb{I}_{\hat{p}^i(\omega_k) \geq p(\omega_i)}}$

Given signal structure π, the receiver's strategy σ^i, upon observing signal realization s_i, is to mix over actions in the set $D = \left\{a_j | \hat{p}^i_j \geq p(\omega_i)\right\}$ in such a way that $\sum_{a \in D} \sigma(a) V_S(a) = V_S(a_{k^*})$.

5. Conclusion

We illustrated the usefulness of reduced form representations for persuasion problems in two ways. In the first, the reduced form reduces the worst-case complexity of determining the optimal persuasion scheme. In the second, it is used to identify a simple algorithm to determine the maximum payoff a sender can achieve in any cheap talk equilibrium.

Declaration of competing interest

None.

Data availability

No data was used for the research described in the article.

Acknowledgments

We thank Elliott Lipnowski, Ilya Segal, and John Quah for valuable suggestions. The research of the first author was supported in part by ERC Grant #757229 under the European Union's Horizon 2020 research and innovation program.

Appendix A. Proof of Theorem 3.1

If V_R satisfies (4), then, $u(p, a) = f(a) + g(a)\mathbb{E}_p\omega + \mathbb{E}_p h$. Thus, if q has the same mean as p, then $u(q, a) = u(p, a) + (\mathbb{E}_q h - \mathbb{E}_p h)$ as desired.

To show the converse, suppose first that there are two states, ω and ω'. Then, every payoff function V_R satisfies (4) with $h \equiv 0$. To see this, fix V_R. Define functions f, g by $f(a) = V_R(\omega, a) - [V_R(\omega', a) - V_R(\omega, a)]\frac{\omega}{\omega' - \omega}$ and $g(a) = [V_R(\omega', a) - V_R(\omega, a)]\frac{1}{\omega' - \omega}$. Then $f(a) + g(a)\omega = V_R(\omega, a)$ and $f(a) + g(a)\omega' = V_R(\omega', a)$. Thus, V_R satisfies (4) with $h \equiv 0$.

Now suppose there are $n \geq 3$ states. Without loss, assume $\omega_1 < \cdots < \omega_n$. By the previous step, the restriction of V_R to the set $\{\omega_1, \omega_n\} \times \mathcal{A}$ satisfies (4) with $h \equiv 0$. That is, there exist f, g such that $V_R(\omega_j, a) = f(a) + g(a)\omega_j$ for all $a \in \mathcal{A}$ for $j \in \{1, n\}$. Let $j \in \{2, \ldots, n-1\}$. Then there exists $\lambda \in (0, 1)$ such that $\lambda\omega_1 + (1-\lambda)\omega_n = \omega_j$. Let p^λ be the corresponding two-point distribution. Since the receiver's preferences depend only on the mean, there exist constants $\alpha_j > 0$ and β_j such that, for all a,

$$\begin{aligned} V_R(\omega_j, a) &= u(\delta_{\omega_j}, a) = \alpha_j u(p^\lambda, a) + \beta_j \\ &= \alpha_j[\lambda V_R(\omega_1, a) + (1-\lambda)V_R(\omega_n, a)] + \beta_j. \end{aligned}$$

Using the form of V_R at ω_1 and ω_n then gives:

$$V_R(\omega_j, a) = \alpha_j[f(a) + g(a)\omega_j] + \beta_j.$$

By inspection of (4), it suffices to show that $\alpha_j = 1$. To this end, take $\mu \in (0, 1)$ and $\eta \in (0, 1)$ such that $0 \neq \bar{\omega} := \mu\omega_1 + (1-\mu)\omega_n = \eta\omega_j + (1-\eta)\omega_n$. Because the receiver's preferences depend only on the posterior mean, there exist constants $\kappa > 0$ and ρ such that, for all a, $u(p^\eta, a) = \kappa u(p^\mu, a) + \rho$, or

$$\eta V_R(\omega_j, a) + (1-\eta)V_R(\omega_n, a) = \kappa[\mu V_R(\omega_1, a) + (1-\mu)V_R(\omega_n, a)] + \rho.$$

Substituting in what we know about V_R gives

$$\eta[\alpha_j(f(a) + g(a)\omega_j) + \beta_j] + (1-\eta)(f(a) + g(a)\omega_n) = \kappa[f(a) + g(a)\bar{\omega}] + \rho.$$

Matching the coefficients of $f(a)$ on both sides gives $\kappa = \eta\alpha_j + 1 - \eta$. Similarly, matching the coefficients of $g(a)$ gives $\kappa = [\eta\alpha_j\omega_j + (1-\eta)\omega_n]/\bar{\omega}$. Thus, the equation can hold for every action a only if these two expressions for κ coincide, which can easily be verified to be the case only if $\alpha_j = 1$.[9]

Appendix B. Proof of Theorem 3.2

The persuasion problem in its obedience formulation assuming quadratic loss for the receiver is:

$$\max \sum_{i=1}^{|\mathcal{A}|} \sum_{j=1}^{|\mathcal{S}|} V_S(a_i) x(\omega_j, a_i) \tag{31}$$

$$\text{s.t.} \sum_{j=1}^{|\mathcal{S}|} [(a_k - \omega_j)^2 - (a_i - \omega_j)^2] x(\omega_j, a_i) \geq 0 \text{ for all } a_i, a_k \in \mathcal{A} \tag{32}$$

$$\sum_{i=1}^{|\mathcal{A}|} x(\omega_j, a_i) = p(\omega_j) \ \forall \omega_j \in \mathcal{S} \tag{33}$$

$$x(\omega_j, a_i) \geq 0 \ \forall a_i \in \mathcal{A}, \omega_j \in \mathcal{S}. \tag{34}$$

Using Eq. (7) the relevant OC are:

$$\frac{a_i + a_{i+1}}{2} \geq \frac{\sum_{j=1}^{|\mathcal{S}|} \omega_j x(\omega_j, a_i)}{\sum_{j=1}^{|\mathcal{S}|} x(\omega_j, a_i)} \geq \frac{a_i + a_{i-1}}{2} \ \forall i \in \{2, \ldots, |\mathcal{A}| - 1\}$$

$$\frac{a_1 + a_2}{2} \geq \frac{\sum_{j=1}^{|\mathcal{S}|} \omega_j x(\omega_j, a_i)}{\sum_{j=1}^{|\mathcal{S}|} x(\omega_j, a_i)}$$

$$\frac{\sum_{j=1}^{|\mathcal{S}|} \omega_j x(\omega_j, a_{|\mathcal{A}|})}{\sum_{j=1}^{|\mathcal{S}|} x(\omega_j, a_{|\mathcal{A}|})} \geq \frac{a_{|\mathcal{A}|} + a_{|\mathcal{A}|-1}}{2}$$

[9] To see this in somewhat more detail, recall that by the maintained assumption in this section, the optimal action for the receiver is not the same in all states. Moreover, we have already shown above that $V_R(\omega_j, a) = \alpha_j(f(a) + g(a)\omega_j) + \beta_j$ for all j. Topkis' Theorem implies that the term $g(a)$ has to be non-decreasing in ω under the optimal action, and hence the optimal actions at ω_1 and at ω_n must be different. (Otherwise, the action that is optimal at these extreme states would be optimal at all states, violating our assumption.) Since $\alpha_1 = \alpha_n = 1$ and $\beta_1 = \beta_n = 0$, this means that there exist actions a and b such that $f(a) + g(a)\omega_1 > f(b) + g(b)\omega_1$ and $f(a) + g(a)\omega_n < f(b) + g(b)\omega_n$. Subtracting the second inequality from the first gives $g(a)(\omega_1 - \omega_n) > g(b)(\omega_1 - \omega_n)$, or $g(b) > g(a)$. But then the first inequality implies $f(a) > f(b)$. Therefore, we have $f(a) \neq f(b)$ and $g(a) \neq g(b)$, and for the equation to hold for both a and b, the two expressions for κ have to coincide.

Using this, we reformulate (32)–(34):

$$-\sum_{j=1}^{|\mathcal{S}|}\omega_j x(\omega_j,a_i)+\sum_{j=1}^{|\mathcal{S}|}\Big(\frac{a_i+a_{i-1}}{2}\Big)x(\omega_j,a_i)\le 0\ \forall i\in\{2,\dots,|\mathcal{A}|\} \tag{35}$$

$$\sum_{j=1}^{|\mathcal{S}|}\omega_j x(\omega_j,a_i)-\sum_{j=1}^{|\mathcal{S}|}\Big(\frac{a_i+a_{i+1}}{2}\Big)x(\omega_j,a_i)\le 0\ \forall i\in\{1,\dots,|\mathcal{A}|-1\} \tag{36}$$

$$\sum_{i=1}^{|\mathcal{A}|}x(\omega_j,a_i)=p(\omega_j)\ \forall\omega_j\in\mathcal{S} \tag{37}$$

$$z_i-\sum_{j=1}^{|\mathcal{S}|}x(\omega_j,a_i)=0\ \forall i\in\{1,\dots,|\mathcal{A}|\} \tag{38}$$

$$x(\omega_j,a_i)\ge 0\ \forall\omega_j\in\mathcal{S}\ a_i\in\mathcal{A}. \tag{39}$$

Our goal is to eliminate the x variables and find an equivalent representation involving just the z variables. Geometrically, we are projecting the polyhedron (35)–(39), which lives in the (x,z) space, into just the z space. We review the basic facts about projections next. For more details see Balas (2001). The reader familiar with this can omit it without loss.

B.1. Projection

Let $P=\{(x,y):Ax+By\le b\}$ where $x\in\mathbb{R}^n$, $y\in\mathbb{R}^k$, $b\in\mathbb{R}^m$, A is a $m\times n$ matrix and B is a $m\times k$ matrix. Assume $P\neq\emptyset$. The **projection** of P into the x space is the set $Q=\{x\in\mathbb{R}^n:\ \exists y\in\mathbb{R}^k \text{st. } (x,y)\in P\}$. We would like to obtain a description of Q. Let $C=\{u\ge 0: uB=0\}$. The set C is a polyhedral cone sometimes called the **elimination cone**. Notice there is one component of u for each inequality in P.

Theorem 7.1.

$Q=\{x: uAx\le ub\ \forall u\in C\}.$

Proof. It is straightforward to see that

$Q\subseteq\{x:uAx\le ub\ u\ge 0\quad uB=0,\quad u\neq 0\}.$

Suppose, for a contradiction that there is an x^* in $\{x: uAx\le ub\quad u\ge 0\quad uB=0,\quad u\neq 0\}$ that is not in Q. This means there is no feasible choice of y in the following system:

$Ax^*+By\le b.$

By the Farkas lemma, there must exist a vector $u\ge 0$ such that $u(b-Ax^*)<0$ and $uB=0$. However, this contradicts the choice of x^*. ■

Let $\mathcal{U}$ be the set of extreme rays of C. An extreme ray is a vector in C that cannot be expressed as non-negative linear combination of other vectors in C. There are a finite number of these. Hence,

$Q=\{x:uAx\le ub\ u\in\mathcal{U}\}.$

If the only solution to $uB=0, u\ge 0$ is the trivial one, then, $Q=\mathbb{R}^n$.

Thus, the problem of characterizing Q reduces to determining the extreme rays of the elimination cone. Identifying the extreme rays of a polyhedral cone is a straightforward but tedious computation involving a variant of Gaussian elimination credited to Fourier and Motzkin (see Khachiyan, 2001). Our goal is not just to compute the extreme rays but find a succinct characterization of them.

Our approach to doing so will be to select an arbitrary $x\in Q$ and focus on $\arg\max\{u(Ax-b):\ \text{s.t. } u\in C\}$. While the feasible region is unbounded (because C is a cone), this linear program has an optimal solution because it is both feasible, and the objective function value is bounded above by zero. The last follows from the fact that $u\in C$. If this program has multiple optima, we can, by scaling, focus on one that satisfies $\mathbf{1}u=1$. In this way, we determine the tangent hyperplanes to Q.

While the polyhedron P in the larger space was described using inequalities only, accommodating equality constraints can be done in the usual way. The component of u corresponding to an equality constraint would be unrestricted in sign.

B.2. The elimination cone

If we set $y(\omega_j,a_i)=\omega_j x(\omega_j,a_i)$, the constraints (35)–(37) can be rewritten as

$$-\sum_{j=1}^{|\mathcal{S}|}y(a_i,\omega_j)+\Big(\frac{a_i+a_{i-1}}{2}\Big)z_i\le 0\ \forall i\in\{2,\dots,|\mathcal{A}|\}\ \ (u_i)$$

$$\sum_{j=1}^{|\mathcal{S}|}y(\omega_j,a_i)-\Big(\frac{a_i+a_{i+1}}{2}\Big)z_i\le 0\ \forall i\in\{1,\dots,|\mathcal{A}|-1\}\ \ (v_i)$$

$$\sum_{i=1}^{|\mathcal{A}|}y(\omega_j,a_i)=\omega_j p(\omega_j)\ \forall\omega_j\in\mathcal{S}\ (w_j)$$

$$z_i-\sum_{j=1}^{|\mathcal{S}|}\omega_j^{-1}y(\omega_j,a_i)=0\ \forall i\in\{1,\dots,|\mathcal{A}|\}\ \ (\lambda_i)$$

We have included with constraint, in parenthesis, the variables that will be used in the description of the elimination cone. The elimination cone is given by

$$-u_i+v_i+w_j-\lambda_i\omega_j^{-1}\ge 0\ \forall 2\le i\le|\mathcal{A}|-1, j\in\mathcal{S} \tag{40}$$

$$-u_{|\mathcal{A}|}+w_j-\lambda_{|\mathcal{A}|}\omega_j^{-1}\ge 0\ \forall j \tag{41}$$

$$v_1+w_j-\lambda_1\omega_j^{-1}\ge 0\ \forall j \tag{42}$$

$$u_i\ge 0\ \forall i\in\{2,\dots,|\mathcal{A}|\} \tag{43}$$

$$v_i\ge 0\ \forall i\in\{1,\dots,|\mathcal{A}|-1\} \tag{44}$$

Each non-trivial element of the elimination cone where at least one of u or v is non-zero gives rise to the following inequality

$$\sum_i\lambda_i z_i+\sum_{i\ge 2}0.5u_i(a_i+a_{i-1})z_i - \sum_{i\le|\mathcal{A}|-1}0.5v_i(a_i+a_{i+1})z_i\le\sum_j w_j\omega_j p(\omega_j). \tag{45}$$

On the other hand, if for all i, $u_i=v_i=0$, $\lambda_i=1$ and $w_j=0$ for all j, we obtain $\sum_i z_i=1$. We assume that the non-zero values of λ are all the same. By scaling we can suppose they are all 1's or all -1's. We justify this at the completion of the proof.

Proposition 7.2. *(x,z) is feasible for (35)–(39) if and only if z is feasible for (15)–(17).*

Proof. The proof is divided into two parts. In the first we suppose the λs are 0-–1 and this will generate (14). In the second part we suppose that $\lambda_i \in \{0, -1\}$ for all $i \in \mathcal{A}$ and this will generate (zloss1).

Part 1: Let $T = \{i : \lambda_i = 1\}$ and focus on elements of the elimination cone where at least one of u or v is non-zero. Choose any feasible z and consider

$$\max \sum_{i\geq 2} 0.5u_i(a_i+a_{i-1})z_i - \sum_{i\leq|\mathcal{A}|-1} 0.5v_i(a_i+a_{i+1})z_i - \sum_j w_j\omega_j p(\omega_j) \tag{46}$$

$$\text{s.t.} \ -u_i + v_i + w_j \geq \omega_j^{-1} \ \forall 2 \leq i \leq |\mathcal{A}|-1, i \in T, j \in \mathcal{S} \tag{47}$$

$$-u_i + v_i + w_j \geq 0 \ \forall 2 \leq i \leq |\mathcal{A}|-1, i \notin T, j \in \mathcal{S} \tag{48}$$

$$-u_{|\mathcal{A}|} + w_j - \lambda_{|\mathcal{A}|}\omega_j^{-1} \geq 0 \ \forall j \tag{49}$$

$$v_1 + w_j - \lambda_1\omega_j^{-1} \geq 0 \ \forall j \tag{50}$$

$$u_i \geq 0 \ \forall i \in \{2, \ldots, |\mathcal{A}|\} \tag{51}$$

$$v_i \geq 0 \ \forall i \in \{1, \ldots, |\mathcal{A}|-1\} \tag{52}$$

Problem (46)–(52) is clearly feasible, and given the choice of z, has a bounded objective function value. Without loss we can assume that $u_iv_i = 0$ for all $i \in \{2, \ldots, |\mathcal{A}|-1\}$. If not, add $\delta < 0$ to both u_i and v_i. Feasibility is preserved and the objective function value changes by $\delta(a_{i-1} - a_i) > 0$ which contradicts optimality.

Choose $K \subseteq \{2, \ldots, |\mathcal{A}|-1\}$ and let K^* be $\{2, \ldots, |\mathcal{A}|-1\} \setminus K$ with at least one of K or K^* being non-empty. We focus on solutions to (46)–(52) where $v_i > 0$ for all $i \in K$ and $u_i \geq 0$ for all $i \in K^*$. The corresponding optimization problem is

$$\max \sum_{i\in K^*\cup|\mathcal{A}|} 0.5u_i(a_i+a_{i-1})z_i - \sum_{i\in K\cup\{1\}} 0.5v_i(a_i+a_{i+1})z_i - \sum_j w_j\omega_j p(\omega_j)$$

$$\text{s.t.} \ v_i + w_j \geq \omega_j^{-1} \ \forall i \in K \cap T, j \in \mathcal{S}$$

$$-u_i + w_j \geq \omega_j^{-1} \ \forall i \in K^* \cap T, j \in \mathcal{S}$$

$$v_i + w_j \geq 0 \ \forall i \in T^c \cap K, j \in \mathcal{S}$$

$$-u_i + w_j \geq 0 \ \forall i \in T^c \cap K^*, j \in \mathcal{S}$$

$$-u_{|\mathcal{A}|} + w_j - \lambda_{|\mathcal{A}|}\omega_j^{-1} \geq 0 \ \forall j$$

$$v_1 + w_j - \lambda_1\omega_j^{-1} \geq 0 \ \forall j$$

$$u_i \geq 0 \ \forall i \in \{2, \ldots, |\mathcal{A}|\}$$

$$v_i \geq 0 \ \forall i \in \{1, \ldots, |\mathcal{A}|-1\}$$

Fixing the value of the w_js, the variables u_i and v_i are bounded as follows:

1. $\forall i \in K \cap T$, $v_i \geq \max_j(\omega_j^{-1} - w_j) = (\omega_{j_1}^{-1} - w_{j_1})$ and $v_i \geq 0$.
2. $\forall i \in K^* \cap T$, $0 \leq u_i \leq \min_j(w_j - \omega_j^{-1}) = w_{j_1} - \omega_{j_1}^{-1}$.
3. $\forall i \in T^c \cap K$, $v_i \geq \max_j -w_j = -\min_j w_j = -w_{j_2}$ and $v_i \geq 0$.
4. $\forall i \in T^c \cap K^*$, $0 \leq u_i \leq \min_j w_j = w_{j_2}$.
5. $v_1 \geq \max_j(\lambda_1\omega_j^{-1} - w_j)$ and $v_1 \geq 0$. Depending on the value of λ the maximum is attained on index j_1 or j_2.
6. $u_{|\mathcal{A}|} \geq \min_j(w_j - \lambda_{|\mathcal{A}|}\omega_j^{-1})$ and $u_{|\mathcal{A}|} \geq 0$. Depending on the value of λ the minimum is attained on index j_1 or j_2.

In an optimal solution, each v_i would be set at its lower bound and each u_i to its upper bound.

Case 1: $\max_j(\omega_j^{-1} - w_j) = (\omega_{j_1}^{-1} - w_{j_1}) \leq 0$.

From item 1 above it follows that $v_i = 0$ for all $i \in K \cap T$. As $w_j \geq \omega_j^{-1} \geq 0$ for all j, from item 3 it follows that $v_i = 0$ for $i \in K \cap T^c$. From item 5 we see that whether we set $\lambda = 1$ or 0 we can always choose $v_1 = 0$. Therefore, our optimization problem becomes

$$\max \sum_{i\in K^*\cup|\mathcal{A}|} 0.5u_i(a_i + a_{i-1})z_i - \sum_j w_j\omega_j p(\omega_j)$$

$$\text{s.t.} \ 0 \leq u_i = w_{j_1} - \omega_{j_1}^{-1} \ \forall i \in K^* \cap T$$

$$0 \leq u_i = w_{j_2} \ \forall i \in K^* \cap T^c$$

$$0 \leq u_{|\mathcal{A}|} = \min_j(w_j - \lambda_{|\mathcal{A}|}\omega_j^{-1})$$

$$\omega_j^{-1} \leq w_j \ \forall j$$

$$w_{j_2} \leq w_j \ \forall j \neq j_2$$

$$w_{j_1} - \omega_{j_1}^{-1} \leq w_j - \omega_j^{-1} \ \forall j \neq j_1$$

Clearly $w_j = \max\{\omega_j^{-1} + w_{j_1} - \omega_{j_1}^{-1}, w_{j_2}\}$ for all $j \neq j_1, j_2$ and the objective function value is piecewise linear in w_{j_1} and w_{j_2}. The only constraints that will be relevant are

$$\omega_{j_2}^{-1} \leq w_{j_2} \tag{53}$$

$$w_{j_2} \leq w_{j_1} \tag{54}$$

$$0 \leq w_{j_1} - \omega_{j_1}^{-1} \leq w_{j_2} - \omega_{j_2}^{-1} \tag{55}$$

The optimal solution must occur where at least one of (53) or (54) binds. If not, we can add δ to w_{j_1} and w_{j_2}, preserve feasibility and change objective function in proportion to δ, contradicting optimality.

If either (53) or (54) binds, we can choose j_1 and j_2 to be the same index, say, index r. Then, the objective function value is a function of w_r alone. The only constraint is $w_r \geq \omega_r^{-1}$ and this must bind otherwise the objective function is unbounded. Hence

$$w_j = \max\{\omega_j^{-1} + w_r - \omega_r^{-1}, w_r\} = \max\{\omega_j^{-1}, \omega_r^{-1}\}$$

for all $j \neq r$. While the optimal choice of r will depend on K^* and T, any choice of r will yield a valid inequality.

The objective function value becomes

$$\omega_r^{-1} \sum_{i\in K^*\cap T^c} 0.5(a_i + a_{i-1})z_i + (\omega_r^{-1} - \lambda_{|\mathcal{A}|}\omega_r^{-1})0.5(a_{|\mathcal{A}|} + a_{|\mathcal{A}|-1})$$
$$- \max\{\omega_j^{-1}, \omega_r^{-1}\} \sum_j \omega_j p(\omega_j)$$

The corresponding inequality is

$$\omega_r \sum_{i\in T} z_i + \sum_{i\in K^*\cap T^c} 0.5(a_i + a_{i-1})z_i + (1 - \lambda_{|\mathcal{A}|})0.5(a_{|\mathcal{A}|} + a_{|\mathcal{A}|-1})$$
$$\leq \max\{\omega_j^{-1}\omega_r, 1\} \sum_j \omega_j p(\omega_j).$$

The strongest version of this inequality for each fixed r occurs when $T = \{i : 0.5(a_i + a_{i-1}) \leq \omega_r\} \cup \{1\}$ (because we were free to choose $\lambda_1 = 1$) and $K^* = T^c$. Therefore, $\lambda_{|\mathcal{A}|} = 0$:

$$\omega_r z_1 + \omega_r \sum_{i \neq 1: 0.5(a_i+a_{i-1}) \leq \omega_r} z_i + \sum_{i: 0.5(a_i+a_{i-1}) > \omega_r} 0.5(a_i + a_{i-1}) z_i \leq \sum_j \max\{\omega_r, \omega_j\} p(\omega_j).$$

Case 2: $\max_j(\omega_j^{-1} - w_j) = (\omega_{j_1}^{-1} - w_{j_1}) \geq 0$ and $w_{j_2} = \min_j w_j \leq 0$.

From item 1 $v_i = \omega_{j_1}^{-1} - w_{j_1}$ for all $i \in K \cap T$. From item 3, $v_i = -w_{j_2}$ for all $i \in K \cap T^c$. From item 2 we have $0 \leq u_i \leq w_{j_1} - \omega_{j_1} \leq 0$ for $i \in K^* \cap T$. Thus, $u_i = 0$ for all $i \in K^* \cap T$ and $w_{j_1} - \omega_{j_1} = 0$. From item 4, $0 \leq u_i = w_{j_2} \leq 0$ for all $j \in K^* \cap T^c$, hence, $u_i = 0 \forall i \in K^* \cap T$ and $w_{j_2} = 0$.

The optimization problem becomes

$$\max \sum_{i \in K^* \cup |\mathcal{A}|} 0.5 u_i (a_i + a_{i-1}) z_i - \sum_{i \in K \cup \{1\}} 0.5 v_i (a_i + a_{i+1}) z_i - \sum_j w_j \omega_j p(\omega_j)$$

$$\text{s.t. } u_i = 0 \ \forall i \in K^*$$

$$0 \leq u_{|\mathcal{A}|} = \min_j (w_j - \lambda_{|\mathcal{A}|} \omega_j^{-1})$$

$$v_i = 0 \ \forall i \in K$$

$$v_1 = \max\{\max_j (\lambda_1 \omega_{j_1}^{-1} - w_{j_1}), 0\}$$

$$\omega_{j_1}^{-1} = w_{j_1}$$

$$0 = w_{j_2} \leq w_j \ \forall j \neq j_2$$

$$0 = w_{j_1} - \omega_{j_1}^{-1} \leq w_j - \omega_j^{-1} \ \forall j \neq j_1$$

Now, whether $\lambda_1 = 1$ or 0, $v_1 = 0$. To maximize, we would set $u_{|\mathcal{A}|}$ as large as possible which we can do by choosing $\lambda_{|\mathcal{A}|} = 0$. Hence, $0 \leq u_{|\mathcal{A}|} = w_{j_2}$, i.e. $u_{|\mathcal{A}|} = 0$. Finally, we set $w_j = \omega_j^{-1}$ for all j. This leaves us with an objective function value of $-\sum_j p(\omega_j) = 1$. The corresponding inequality is $\sum_{i \in T} z_i - 1 \leq 0$. The strongest version of this is $\sum_{i \in \mathcal{A}} z_i \leq 1$.

Case 3: $\max_j(\omega_j^{-1} - w_j) = (\omega_{j_1}^{-1} - w_{j_1}) \geq 0$ and $w_{j_2} = \min_j w_j \geq 0$.

From item 1 $v_i = \omega_{j_1}^{-1} - w_{j_1}$ for all $i \in K \cap T$. From item 3, $v_i = \max\{-w_{j_2}, 0\} = 0$ for all $i \in K \cap T^c$. From item 2 we have $0 \leq u_i \leq w_{j_1} - \omega_{j_1} \leq 0$ for $i \in K^* \cap T$. Thus, $u_i = 0$ for all $i \in K^* \cap T$ and $w_{j_1} - \omega_{j_1} = 0$. From item 4, $0 \leq u_i = w_{j_2}$ for all $j \in K^* \cap T^c$,

The optimization problem is

$$\max \sum_{i \in K^* \cup |\mathcal{A}|} 0.5 u_i (a_i + a_{i-1}) z_i - \sum_{i \in K \cup \{1\}} 0.5 v_i (a_i + a_{i+1}) z_i - \sum_j w_j \omega_j p(\omega_j)$$

$$\text{s.t. } u_i = 0 \ \forall i \in K^* \cap T$$

$$u_i = w_{j_2} \ \forall i \in K^* \cap T^c$$

$$u_{|\mathcal{A}|} = \min_j (w_j - \lambda_{|\mathcal{A}|} \omega_j^{-1})$$

$$v_i = 0 \ \forall i \in K$$

$$v_1 = \max\{\max_j (\lambda_1 \omega_{j_1}^{-1} - w_{j_1}), 0\}$$

$$\omega_{j_1}^{-1} = w_{j_1}$$

$$0 \leq w_{j_2} \leq w_j \ \forall j \neq j_2$$

$$w_{j_1} - \omega_{j_1}^{-1} \leq w_j - \omega_j^{-1} \ \forall j \neq j_1$$

To optimize, we would set $\lambda_{|\mathcal{A}|} = 0$ and $\lambda_1 = 0$. Hence $u_{|\mathcal{A}|} = w_{j_2}$ and $v_1 = 0$.

Our optimization problem reduces to

$$\max \sum_{i \in [K^* \cap T^c] \cup |\mathcal{A}|} 0.5 w_{j_2} (a_i + a_{i-1}) z_i - \sum_j w_j \omega_j p(\omega_j)$$

$$\text{s.t. } 0 \leq w_{j_2} \leq w_j \ \forall j \neq j_2$$

$$0 \leq w_j - \omega_j^{-1} \ \forall j \neq j_1$$

Therefore, at optimality $w_j = \max\{w_{j_2}, \omega_j^{-1}\}$. The optimal solution must occur at some breakpoint, say $w_{j_2} = \omega_r^{-1}$. The corresponding inequality is

$$\sum_{i \in T} z_i + \omega_r^{-1} \sum_{i \in [K^* \cap T^c] \cup |\mathcal{A}|} 0.5(a_i + a_{i-1}) z_i \leq \sum_j \max\{\omega_r^{-1}, \omega_j^{-1}\} \omega_j p(\omega_j)$$

$$\omega_r \sum_{i \in T} z_i + \sum_{i \in [K^* \cap T^c] \cup |\mathcal{A}|} 0.5(a_i + a_{i-1}) z_i \leq \sum_j \max\{\omega_r, \omega_j\} p(\omega_j)$$

However, this is the same inequality we had in case 1.

Part 2: Now, let $T = \{i : \lambda_i = -1\}$. As before, choose any feasible z and consider

$$\max \sum_{i \geq 2} 0.5 u_i (a_i + a_{i-1}) z_i - \sum_{i \leq |\mathcal{A}|-1} 0.5 v_i (a_i + a_{i+1}) z_i - \sum_j w_j \omega_j p(\omega_j) \tag{56}$$

$$\text{s.t. } -u_i + v_i + w_j \geq -\omega_j^{-1} \ \forall 2 \leq i \leq |\mathcal{A}| - 1, i \in T, j \in \mathcal{S} \tag{57}$$

$$-u_i + v_i + w_j \geq 0 \ \forall 2 \leq i \leq |\mathcal{A}| - 1, i \notin T, j \in \mathcal{S} \tag{58}$$

$$-u_{|\mathcal{A}|} + w_j - \lambda_{|\mathcal{A}|} \omega_j^{-1} \geq 0 \ \forall j \tag{59}$$

$$v_1 + w_j - \lambda_1 \omega_j^{-1} \geq 0 \ \forall j \tag{60}$$

$$u_i \geq 0 \ \forall i \in \{2, \ldots, |\mathcal{A}|\} \tag{61}$$

$$v_i \geq 0 \ \forall i \in \{1, \ldots, |\mathcal{A}| - 1\} \tag{62}$$

Problem (56)–(62) is feasible and has a bounded objective function value. As before we can assume that $u_i v_i = 0$ for all $i \in \{2, \ldots, |\mathcal{A}| - 1\}$.

Choose $K \subseteq \{2, \ldots, |\mathcal{A}| - 1\}$ and let K^* be $\{2, \ldots, |\mathcal{A}| - 1\} \setminus K$ with at least one of K or K^* being non-empty. We focus on solutions to (56)–(62) where $v_i > 0$ for all $i \in K$ and $u_i \geq 0$ for all $i \in K^*$. The corresponding optimization problem is

$$\max \sum_{i \in K^* \cup |\mathcal{A}|} 0.5 u_i (a_i + a_{i-1}) z_i - \sum_{i \in K \cup \{1\}} 0.5 v_i (a_i + a_{i+1}) z_i - \sum_j w_j \omega_j p(\omega_j)$$

$$\text{s.t. } v_i + w_j \geq -\omega_j^{-1} \ \forall i \in K \cap T, j \in \mathcal{S}$$

$$-u_i + w_j \geq -\omega_j^{-1} \ \forall i \in K^* \cap T, j \in \mathcal{S}$$

$$v_i + w_j \geq 0 \ \forall i \in T^c \cap K, j \in \mathcal{S}$$

$$-u_i + w_j \geq 0 \ \forall i \in T^c \cap K^*, j \in \mathcal{S}$$

$$-u_{|\mathcal{A}|} + w_j - \lambda_{|\mathcal{A}|}\omega_j^{-1} \geq 0 \ \forall j$$

$$v_1 + w_j - \lambda_1\omega_j^{-1} \geq 0 \ \forall j$$

$$u_i \geq 0 \ \forall i \in \{2, \ldots, |\mathcal{A}|\}$$

$$v_i \geq 0 \ \forall i \in \{1, \ldots, |\mathcal{A}| - 1\}$$

Fixing the value of the w_js (these can be negative), the variables u_i and v_i are determined as follows:

1. $v_i \geq \max\{\max_j(-\omega_j^{-1} - w_j), 0\} \ \forall i \in K \cap T$
2. $0 \leq u_i \leq \min_j(w_j + \omega_j^{-1}) \ \forall i \in K^* \cap T$.
3. $v_i \geq \max\{\max_j -w_j, 0\} \ \forall i \in K \cap T^c$
4. $0 \leq u_i \leq \min_j w_j \ \forall i \in K^* \cap T^c$
5. $v_1 \geq \max\{\max_j(\lambda_1\omega_j^{-1} - w_j), 0\}$
6. $0 \leq u_{|\mathcal{A}|} \leq \min_j(w_j - \lambda_{|\mathcal{A}|}\omega_j^{-1})$

Case 1: $\max_j(-\omega_j^{-1} - w_j) = -\omega_{j_1}^{-1} - w_{j_1} \geq 0$.

By item 1 above $v_i = -\omega_{j_1}^{-1} - w_{j_1} \geq 0$ for all $i \in K \cap T$. By item 2 above we have

$$0 \leq u_i \leq \min_j(w_j + \omega_j^{-1}) \leq 0.$$

Therefore $u_i = 0$ for all $i \in K^* \cap T$ and $(w_{j_1} + \omega_{j_1}^{-1}) = 0$. Hence, $v_i = 0$ for all $i \in K \cap T$.

Now, $w_{j_1} = -\omega_{j_1}^{-1} \leq 0$ implies that $\min_j w_j \leq 0$. Therefore, by item 3 above $v_i = -\min_j w_j = -w_{j_2} \geq 0$ for all $i \in K \cap T^c$.

By item 4 above, $0 \leq u_i \leq \min_j w_j \leq 0$ for all $i \in T^c \cap K^*$. Either $\min_j w_j = 0$ or $T^c \cap K^* = \emptyset$.

In the first case $u_i = 0$ for all $i \in K^*$, $v_i = 0$ for all $i \in K$ and the optimization problem becomes

$$\max 0.5u_{|\mathcal{A}|}(a_{|\mathcal{A}|} + a_{|\mathcal{A}|-1})z_{|\mathcal{A}|} - 0.5v_1(a_1 + a_2)z_1 - \sum_j w_j\omega_j p(\omega_j)$$

$$\text{s.t. } 0 \leq u_{|\mathcal{A}|} = \min_j(w_j - \lambda_{|\mathcal{A}|}\omega_j^{-1})$$

$$0 \leq v_1 = \max_j(\lambda_1\omega_j^{-1} - w_j)$$

$$\omega_{j_1}^{-1} + w_{j_1} = 0$$

$$-w_{j_2} \geq -w_j \ \forall j \neq j_2$$

$$0 \leq w_j + \omega_j^{-1} \ \forall j \neq j_1$$

If $\lambda_{|\mathcal{A}|} = 0$, then $0 \leq u_{|\mathcal{A}|} \leq \min_j w_j = 0$. In that case we would set each w_j as small as possible which is $\max\{0, -\omega_j^{-1}\} = 0$. This gives rise to the trivial inequality $\sum_{i \in T} z_i \geq 0$. If $\lambda_{|\mathcal{A}|} = -1$, then $0 \leq u_{|\mathcal{A}|} \min_j(w_j + \omega_j^{-1}) \leq 0$. Again, we obtain a trivial inequality.

So we go on to consider the next possibility, which means that $K^* \subseteq T$. The optimization problem becomes

$$\max \sum_{i \in K^* \cup |\mathcal{A}|} 0.5u_i(a_i + a_{i-1})z_i - \sum_{i \in K \cup \{1\}} 0.5v_i(a_i + a_{i+1})z_i - \sum_j w_j\omega_j p(\omega_j)$$

$$\text{s.t. } u_i = 0 \ \forall i \in K^*$$

$$0 \leq u_{|\mathcal{A}|} = \min_j(w_j - \lambda_{|\mathcal{A}|}\omega_j^{-1})$$

$$0 \leq v_i = -w_{j_2} \ \forall i \in K \cap T^c$$

$$v_i = 0 \ \forall i \in K \cap T$$

$$v_1 = \max\{\max_j(\lambda_1\omega_j^{-1} - w_j), 0\}$$

$$\omega_{j_1}^{-1} + w_{j_1} = 0$$

$$-w_{j_2} \geq -w_j \ \forall j \neq j_2$$

$$0 \leq w_j + \omega_j^{-1} \ \forall j \neq j_1$$

Whether we set $\lambda_1 = -1$ or $\lambda_1 = 0$, v_1 is always zero. Feasibility requires that $\lambda_{|\mathcal{A}|} = -1$ which forces $u_{|\mathcal{A}|} = 0$. So, our optimization problem becomes:

$$\max w_{j_2} \sum_{i \in K \cap T^c} 0.5(a_i + a_{i+1})z_i - \sum_j w_j\omega_j p(\omega_j)$$

$$\text{s.t. } \omega_{j_1}^{-1} + w_{j_1} = 0$$

$$-w_{j_2} \geq -w_j \ \forall j \neq j_2$$

$$0 \leq w_j + \omega_j^{-1} \ \forall j \neq j_1$$

$$w_{j_2} \leq 0$$

The constraints reduce to $-\omega_{j_2}^{-1} \leq w_{j_2} \leq -\omega_{j_1}^{-1}$ and $w_j = \max\{w_{j_2}, -\omega_j^{-1}\}$ for all $j \neq j_2$. So, we can write the optimization problem as

$$\max w_{j_2} \sum_{i \in K \cap T^c} 0.5(a_i + a_{i+1})z_i - \sum_{j \neq j_2} \max\{w_{j_2}, -\omega_j^{-1}\}\omega_j p(\omega_j) - w_{j_2}\omega_{j_2}p(\omega_{j_2})$$

$$\text{s.t. } -\omega_{j_2}^{-1} \leq w_{j_2} \leq -\omega_{j_1}^{-1}$$

At optimality w_{j_2} must be at its upper or lower bound. Suppose first that $-\omega_{j_2}^{-1} = w_{j_2}$. The objective function value becomes

$$-\omega_{j_2}^{-1} \sum_{i \in K \cap T^c} 0.5(a_i + a_{i+1})z_i + p(\omega_{j_2}) - \sum_{j \neq j_2} \max\{-\omega_{j_2}^{-1}, -\omega_j^{-1}\}\omega_j p(\omega_j)$$

The corresponding inequality is

$$-\sum_{i \in T} z_i - \omega_{j_2}^{-1} \sum_{i \in K \cap T^c} 0.5(a_i + a_{i+1})z_i + p(\omega_{j_2}) - \sum_{j \neq j_2} \max\{-\omega_{j_2}^{-1}, -\omega_j^{-1}\}\omega_j p(\omega_j) \geq 0$$

$$\omega_{j_2} \sum_{i \in T} z_i + \sum_{i \in K \cap T^c} 0.5(a_i + a_{i+1})z_i \geq \omega_{j_2}p(\omega_{j_2}) - \sum_{j \neq j_2} \max\{-1, -\omega_{j_2}\omega_j^{-1}\}\omega_j p(\omega_j)$$

$$\omega_{j_2} \sum_{i \in T} z_i + \sum_{i \in K \cap T^c} 0.5(a_i + a_{i+1})z_i \geq \sum_j \min\{\omega_j, \omega_{j_2}\}p(\omega_j)$$

Because $K^* \subseteq T$ it means that $K = T^c$ and so the strongest version of this inequality occurs when $T = \{i : 0.5(a_i + a_{i+1}) > \omega_{j_2}\} \cup \{|\mathcal{A}|\}$.

$$\omega_{j_2} z_{|\mathcal{A}|} + \omega_{j_2} \sum_{i \neq |\mathcal{A}|: 0.5(a_i+a_{i+1}) > \omega_{j_2}} z_i + \sum_{i: 0.5(a_i+a_{i+1}) \leq \omega_{j_2}} 0.5(a_i + a_{i+1}) z_i \geq \sum_j \min\{\omega_j, \omega_{j_2}\} p(\omega_j)$$

The second possibility is that $w_{j_2} = -\omega_{j_1}^{-1}$. The objective function becomes:

$$-\omega_{j_1}^{-1} \sum_{i \in K \cap T^c} 0.5(a_i + a_{i+1}) z_i - \sum_j \max\{-\omega_{j_1}^{-1}, -\omega_j^{-1}\} \omega_j p(\omega_j)$$

But this yields the same inequality as before.

Case 2: $(w_{j_1} + \omega_{j_1}^{-1}) = \min_j (w_j + \omega_j^{-1}) \geq 0$ and $w_{j_2} = \min_j w_j \geq 0$.

By item 1 we have that $v_i = 0$ for all $i \in K \cap T$. By item 3 we have that $v_i = 0$ for all $i \in K \cap T^c$. By item 3, $u_i = w_{j_1} + \omega_{j_1}^{-1}$ for all $i \in K^* \cap T$. By item 4, $u_i = w_{j_2}$ for all $i \in K^* \cap T^c$.

The optimization problem becomes

$$\max \sum_{i \in K^* \cup |\mathcal{A}|} 0.5 u_i (a_i + a_{i-1}) z_i - \sum_{i \in K \cup \{1\}} 0.5 v_i (a_i + a_{i+1}) z_i - \sum_j w_j \omega_j p(\omega_j)$$

$$\text{s.t. } v_i = 0 \forall i \in K$$

$$u_i = w_{j_1} + \omega_{j_1}^{-1} \ \forall i \in K^* \cap T$$

$$u_i = w_{j_2} \ \forall i \in T^c \cap K^*$$

$$u_{|\mathcal{A}|} = \min_j (w_j - \lambda_{|\mathcal{A}|} \omega_j^{-1})$$

$$v_1 \geq \max_j (\lambda_1 \omega_j^{-1} - w_j)$$

$$\omega_{j_1}^{-1} + w_{j_1} \geq 0$$

$$w_{j_2} \leq w_j \ \forall j \neq j_2$$

$$w_{j_1} + \omega_{j_1}^{-1} \leq w_j + \omega_j^{-1} \ \forall j \neq j_1$$

$$v_1 \geq 0$$

Whether we set $\lambda_1 = -1$ or zero, we would still set $v_1 = 0$. The problem becomes

$$\max \sum_{i \in K^* \cup |\mathcal{A}|} 0.5 u_i (a_i + a_{i-1}) z_i - \sum_j w_j \omega_j p(\omega_j)$$

$$\text{s.t. } u_i = w_{j_1} + \omega_{j_1}^{-1} \ \forall i \in K^* \cap T$$

$$u_i = w_{j_2} \ \forall i \in T^c \cap K^*$$

$$u_{|\mathcal{A}|} = \min_j (w_j - \lambda_{|\mathcal{A}|} \omega_j^{-1})$$

$$\omega_{j_1}^{-1} + w_{j_1} \geq 0$$

$$w_{j_2} \leq w_j \ \forall j \neq j_2$$

$$w_{j_1} + \omega_{j_1}^{-1} \leq w_j + \omega_j^{-1} \ \forall j \neq j_1$$

Observe, if we add δ to all w_j, feasibility is preserved and the objective function changes linearly in δ. If objective function value increases with δ this would violate boundedness. So, it must be that objective function value increases with $\delta < 0$. Therefore, we would set $\delta = -w_{j_2}$, meaning that in our solution $w_{j_2} = 0$. The constraints of our problem reduce to

$$u_i = w_{j_1} + \omega_{j_1}^{-1} \ \forall i \in K^* \cap T$$

$$u_i = 0 \ \forall i \in T^c \cap K^*$$

$$u_{|\mathcal{A}|} = \min_j (w_j - \lambda_{|\mathcal{A}|} \omega_j^{-1})$$

$$\omega_{j_1}^{-1} + w_{j_1} \geq 0$$

$$0 \leq w_j \ \forall j \neq j_2$$

$$w_{j_1} + \omega_{j_1}^{-1} \leq w_j + \omega_j^{-1} \ \forall j \neq j_1, j_2$$

To optimize we set $w_j = w_{j_1} + \omega_{j_1}^{-1} - \omega_j^{-1}$ for all $j \neq j_1, j_2$. Suppose $\lambda_{|\mathcal{A}|} = 1$. Then, $u_{|\mathcal{A}|} = u_{|\mathcal{A}|} = \min_j (w_j + \omega_j^{-1}) = w_{j_1} + \omega_{j_1}^{-1}$ and the optimization problem becomes

$$\max \sum_{i \in T \cap K^* \cup \{|\mathcal{A}|\}} 0.5 (w_{j_1} + \omega_{j_1}^{-1})(a_i + a_{i-1}) z_i - \sum_j [w_{j_1} + \omega_{j_1}^{-1} - \omega_j^{-1}] \omega_j p(\omega_j)$$

$$\text{s.t. } \omega_j^{-1} \leq w_{j_1} + \omega_{j_1}^{-1} \leq \omega_{j_2}^{-1} \ \forall j$$

$$w_{j_1} \geq 0$$

Feasibility requires that

$$\omega_1^{-1} \leq \omega_{j_2}^{-1} \leq \omega_1^{-1}.$$

Hence, $w_{j_1} = 0$ and $\omega_{j_1} = \omega_1$. The objective function value is

$$\sum_{i \in T \cap K^* \cup \{|\mathcal{A}|\}} 0.5 \omega_1^{-1} (a_i + a_{i-1}) z_i - \sum_j [\omega_1^{-1} - \omega_j^{-1}] \omega_j p(\omega_j)$$

The corresponding inequality is

$$-\sum_{i \in T} z_i + \sum_{i \in T \cap K^* \cup \{|\mathcal{A}|\}} 0.5 \omega_1^{-1} (a_i + a_{i-1}) z_i - \sum_j [\omega_1^{-1} - \omega_j^{-1}] \omega_j p(\omega_j) \leq 0$$

$$-\omega_1 \sum_{i \in T} z_i + \sum_{i \in T \cap K^* \cup \{|\mathcal{A}|\}} 0.5 (a_i + a_{i-1}) z_i \leq \sum_j [1 - \omega_1 \omega_j^{-1}] \omega_j p(\omega_j)$$

$$-\omega_1 \sum_{i \in T} z_i + \sum_{i \in T \cap K^* \cup \{|\mathcal{A}|\}} 0.5 (a_i + a_{i-1}) z_i \leq \sum_j \omega_j p(\omega_j) - \omega_1$$

$$\omega_1 [1 - \sum_{i \in T} z_i] + \sum_{i \in T \cap K^* \cup \{|\mathcal{A}|\}} 0.5 (a_i + a_{i-1}) z_i \leq \sum_j \omega_j p(\omega_j)$$

The strongest version of this is when $T = K^* \cup \{|\mathcal{A}|\}$ and $K^* = \{i : 0.5(a_i + a_{i-1}) \geq \omega_1, \ 2 \leq i \leq |\mathcal{A}| - 1\}$. The inequality becomes

$$\omega_1 z_1 + \sum_{i=2}^{|\mathcal{A}|} 0.5 (a_i + a_{i-1}) z_i \leq \sum_j \omega_j p(\omega_j).$$

Had we set $\lambda_{|\mathcal{A}|} = 0$ instead, we obtain the weaker inequality:

$$\omega_1 z_1 + \sum_{i=2}^{|\mathcal{A}|-1} 0.5 (a_i + a_{i-1}) z_i \leq \sum_j \omega_j p(\omega_j)$$

Case 3: $(w_{j_1} + \omega_{j_1}^{-1}) = \min_j (w_j + \omega_j^{-1}) \geq 0$ and $w_{j_2} = \min_j w_j \leq 0$.

By item 1 $v_i = 0$ for all $i \in K \cap T$. By item 2, $u_i = (w_{j_1} + \omega_{j_1}^{-1})$ for all $i \in K^* \cap T$. By item 3, $v_i = -w_{j_2}$ for all $i \in K \cap T^c$. Item 4 implies that $w_{j_2} = 0$ and $u_i = 0$ for all $i \in K^* \cap T^c$. Hence, $v_i = 0$ for all $i \in K \cap T^c$.

The optimization problem is

$$\max \sum_{i \in K^* \cup |\mathcal{A}|} 0.5u_i(a_i + a_{i-1})z_i - \sum_{i \in K \cup \{1\}} 0.5v_i(a_i + a_{i+1})z_i - \sum_j w_j \omega_j p(\omega_j)$$

$$\text{s.t. } u_i = w_{j_1} + \omega_{j_1}^{-1} \ \forall i \in K^* \cap T$$

$$u_i = 0 \ \forall i \in K^* \cap T^c$$

$$0 \leq u_{|\mathcal{A}|} = \min_j (w_j - \lambda_{|\mathcal{A}|}\omega_j^{-1})$$

$$v_i = 0 \ \forall i \in K$$

$$v_1 \geq \max\{\max_j(\lambda_1 \omega_j^{-1} - w_j), 0\}$$

$$0 \leq w_{j_1} + \omega_{j_1}^{-1} \leq w_j + \omega_j^{-1} \ \forall j \neq j_1$$

$$w_j \geq w_{j_2} = 0 \ \forall j$$

To optimize we set $\lambda_{|\mathcal{A}|} = -1$ and $\lambda_1 = 0$. Hence, $u_{|\mathcal{A}|} = w_{j_1} + \omega_{j_1}^{-1}$ and $v_1 = 0$. The optimization problem reduces to

$$\max(w_{j_1} + \omega_{j_1}^{-1}) \sum_{i \in K^* \cap T \cup \{|\mathcal{A}|\}} 0.5(a_i + a_{i-1})z_i - \sum_j w_j \omega_j p(\omega_j)$$

$$\text{s.t. } w_{j_1} + \omega_{j_1}^{-1} \leq w_j + \omega_j^{-1} \ \forall j \neq j_1$$

$$w_j \geq w_{j_2} = 0 \ \forall j$$

The optimal objective function value is

$$(w_{j_1} + \omega_{j_1}^{-1}) \sum_{i \in K^* \cap T \cup \{|\mathcal{A}|\}} 0.5(a_i + a_{i-1})z_i - \sum_j \max\{w_{j_1} + \omega_{j_1}^{-1} - \omega_j^{-1}, 0\}\omega_j p(\omega_j)$$

This is piecewise linear in w_{j_1} and the optimal must occur at a breakpoint. Hence, there is an $r \in \mathcal{S}$ such that $w_{j_1} = \omega_r^{-1} - \omega_{j_1}^{1} \geq 0$.

The corresponding inequality is

$$-\sum_{i \in T} z_i + \omega_r^{-1} \sum_{i \in K^* \cap T \cup \{|\mathcal{A}|\}} 0.5(a_i + a_{i-1})z_i - \sum_j \max\{\omega_r^{-1} - \omega_j^{-1}, 0\}\omega_j p(\omega_j) \leq 0$$

$$-\omega_r \sum_{i \in T} z_i + \sum_{i \in K^* \cap T \cup \{|\mathcal{A}|\}} 0.5(a_i + a_{i-1})z_i - \sum_j \max\{1 - \omega_r \omega_j^{-1}, 0\}\omega_j p(\omega_j) \leq 0$$

$$-\omega_r \sum_{i \in T} z_i + \sum_{i \in K^* \cap T \cup \{|\mathcal{A}|\}} 0.5(a_i + a_{i-1})z_i \leq \sum_j \max\{\omega_j - \omega_r, 0\}p(\omega_j)$$

The strongest version of the inequality is when $T = K^* \cup \{|\mathcal{A}|\}$ and $K^* = \{i : 0.5(a_i + a_{i-1}) \geq \omega_r, \ 2 \leq i \leq |\mathcal{A}| - 1\}$.

$$\omega_r \sum_{i \notin T} z_i + \sum_{i \in T} 0.5(a_i + a_{i-1})z_i \leq \sum_j \max\{\omega_j - \omega_r, 0\}p(\omega_j) + \omega_r$$

$$\omega_r \sum_{i: 0.5(a_i + a_{i-1}) < \omega_r} z_i + \sum_{i 0.5(a_i + a_{i-1}) \geq \omega_r} 0.5(a_i + a_{i-1})z_i \leq \sum_j \max\{\omega_j - \omega_r, 0\}p(\omega_j) + \omega_r$$

The right hand side satisfies:

$$\sum_j \max\{\omega_j - \omega_r, 0\}p(\omega_j) + \omega_r = \sum_{j \geq r}(\omega_j - \omega_r)p(\omega_j) + \omega_r \sum_j p(\omega_j) = \sum_j \max\{\omega_r, \omega_j\}p(\omega_j) \quad \blacksquare$$

We now justify why the non-zero components of λ can be chosen to be equal. The dual to (46)–(52) is

$$-\max \sum_i \sum_j \lambda_i \omega_j^{-1} \alpha_{ij}$$

$$\text{s.t. } \sum_j \alpha_{ij} \geq 0.5(a_i + a_{i+1})z_i \ \forall i \in K \cup \{1\}$$

$$\sum_j \alpha_{ij} \leq 0.5(a_i + a_{i-1})z_i \ \forall i \in K^* \cup \{|\mathcal{A}|\}$$

$$\sum_i \alpha_{ij} = \omega_j p(\omega_j)$$

This is an instance of a factored transportation problem (see Evans, 1984), so the solution is 'assortative' in that one pairs high λ with high ω^{-1} and sends as much flow as possible along that arc. Therefore, the optimal solution to the dual is *independent* of the magnitude of the λs; it only depends on how they are ordered from largest to smallest. If we are free to choose the λs to make the objective function value of the primal as large as possible, i.e., the dual (without the negative sign) as small as possible, we could shift weight from large λs to small ones without changing the ordering of the λs. Thus, we can assume that either each λ is zero or, when non-zero, are all equal. Hence, by scaling, we can assume the non-zero entries are all 1 or all -1.

Appendix C. Senders' preferences depend only on posterior mean

The sender's payoff at action i is:

$$\phi(a_i)\frac{\sum_{\omega_j \in \mathcal{S}} \omega_j x(\omega_j, a_i)}{\sum_{j \in \mathcal{S}} x(\omega_j, a_i)}$$

where $x(\omega_j, a_i)$ has the usual meaning.

Thus, the persuasion problem (6)–(7) reduces to:

$$\max \sum_{a_i} \sum_{j \in \mathcal{S}} x(\omega_j, a_i)\phi(a_i)\frac{\sum_{\omega_j \in \mathcal{S}} \omega_j x(\omega_j, a_i)}{\sum_{\omega_j \in \mathcal{S}} x(\omega_j, a_i)}$$

$$\text{s.t. } \frac{f(a_{i_U}) - f(a_i)}{g(a_i) - g(a_{i_U})} \geq \frac{\sum_{\omega_j \in \mathcal{S}} \omega_j x(\omega_j, a_i)}{\sum_{\omega_j \in \mathcal{S}} x(\omega_j, a_i)} \geq \frac{f(a_{i_B}) - f(a_i)}{g(a_i) - g(a_{i_B})} \ \forall a_i$$

For convenience set $U_i = \frac{f(a_{i_U}) - f(a_i)}{g(a_i) - g(a_{i_U})} \geq 0$ and $L_i = \frac{f(a_{i_B}) - f(a_i)}{g(a_i) - g(a_{i_B})} \geq 0$ for all i. Assume that $U_i, L_i \neq 0$ for all i. If we set $y_i = \sum_{\omega_j \in \mathcal{S}} \omega_j x(\omega_j, a_i)$, the sender's optimization problem becomes:

$$\max \sum_{a_i \in \mathcal{A}} \phi(a_i) y_i \tag{63}$$

$$\text{s.t. } y_i - U_i \sum_{\omega_j \in \mathcal{S}} x(\omega_j, a_i) \leq 0 \ \forall a_i \in \mathcal{A} \tag{64}$$

$$-y_i + L_i \sum_{\omega_j \in \mathcal{S}} x(\omega_j, a_i) \leq 0 \ \forall a_i \in \mathcal{A} \tag{65}$$

$$\sum_{a_i} x(\omega_j, a_i) = p(j) \ \forall \omega_j \in \mathcal{S} \tag{66}$$

$$y_i - \sum_{\omega_j \in \mathcal{S}} \omega_j x(\omega_j, a_i) = 0 \ \forall a_i \in \mathcal{A} \tag{67}$$

$$x(\omega_j, a_i), y_i \geq 0 \ \forall a_i \in \mathcal{A}, \omega_j \in \mathcal{S} \tag{68}$$

Theorem 8.1. *Problem (63)–(68) is equivalent to:*

$$\max \sum_{a_i \in \mathcal{A}} \phi(a_i) y_i$$

$$s.t. \sum_{i \in \mathcal{A}} y_i = \sum_{\omega_j \in \mathcal{S}} \omega_j p(j).$$

$$\sum_{i \in \mathcal{A}: \omega_r \geq U_i} [\frac{\omega_r}{U_i} - 1] y_i \leq \sum_{\omega_j : \omega_j \leq \omega_r} (\omega_r - \omega_j) p(j) \ \forall \omega_r \in \mathcal{S}$$

$$\sum_{a_i : \omega_r \leq L_i} [1 - \frac{\omega_r}{L_i}] y_i \leq \sum_{\omega_j : \omega_j \geq \omega_r} (\omega_j - \omega_r) p(j) \ \forall \omega_r \in \mathcal{S}$$

$$y_i \geq 0 \ \forall a_i \in \mathcal{A}$$

The proof is similar to the proof of Theorem 3.2 and is omitted (but is available upon request from the authors).

References

Balas, E., 2001. Projection and lifting in combinatorial optimization. In: Jünger, M., Naddef, D. (Eds.), Computational Combinatorial Optimization. In: Lecture Notes in Computer Science, vol. 2241, Springer, pp. 26–56.
Bergemann, D., Bonatti, A., Smolin, A., 2018. The design and price of information. Amer. Econ. Rev. 108 (1), 1–48.
Bergemann, D., Morris, S., 2016. Bayes correlated equilibrium and the comparison of information structures in games. Theor. Econ. 11 (2), 487–522.
Bertsimas, D., Niño-Mora, J., 1996. Conservation laws, extended polymatroids and multiarmed bandit problems; A polyhedral approach to indexable systems. Math. Oper. Res. 21, 257–306.
Brocas, I., Carrillo, J.D., 2007. Influence through ignorance. Rand J. Econ. 38 (4), 931–947.
Che, Y.-K., Kim, J., Mierendorff, K., 2013. Generalized reduced-form auctions: A network-flow approach. Econometrica 81 (6), 2487–2520.
Cohen, M.B., Lee, Y.T., Song, Z., 2019. Solving linear programs in the current matrix multiplication time. In: Proceedings of the 51st Annual ACM SIGACT Symposium on Theory of Computing. STOC 2019, Association for Computing Machinery, New York, NY, USA, pp. 938–942.
Doval, L., Skreta, V., 2021. Constrained Information Design. Working Paper.
Dughmi, S., Niazadeh, R., Psomas, A., Weinberg, S.M., 2019. Persuasion and incentives through the lens of duality. In: Caragiannis, I., Mirrokni, V., Nikolova, E. (Eds.), Web and Internet Economics. Springer International Publishing, Cham, pp. 142–155.
Dughmi, S., Xu, H., 2016. Algorithmic Bayesian Persuasion. Technical Report.
Dworczak, P., Martini, G., 2019. The simple economics of optimal persuasion. J. Polit. Econ. 127 (5), 1993–2048.
Epitropou, M., Vohra, R.V., 2019. Optimal on-line allocation rules with verification. In: SAGT.
Evans, J.R., 1984. The factored transportation problem. Manage. Sci. 30 (8), 1021–1024.
Gale, D., 1957. A theorem on flows in networks. Pacific J. Math. 7 (2), 1073–1082.
Galperti, S., Perego, J., 2018. A Dual Perspective on Information Design. Technical Report.
Gentzkow, M., Kamenica, E., 2016. A Rothschild-Stiglitz approach to Bayesian persuasion. Am. Econ. Rev. 106 (5), 597–601.
Kamenica, E., 2019. Bayesian persuasion and information design. Annu. Rev. Econ. 11 (1), 249–272.
Kamenica, E., Gentzkow, M., 2011. Bayesian persuasion. Amer. Econ. Rev. 101 (6), 2590–2615.
Khachiyan, L., 2001. Fourier-Motzkin elimination method. In: Encyclopedia of Optimization. Springer US, Boston, MA, pp. 695–699.
Kolotilin, A., 2018. Optimal information disclosure: A linear programming approach. Theor. Econ. 13, 607–635.
Lipnowski, E., Mathevet, L., 2018. Disclosure to a psychological audience. Am. Econ. J. Microecon. 10 (4), 67–93.
Lipnowski, E., Ravid, D., 2020. Cheap talk with transparent motives. Econometrica 88 (4), 1631–1660.
Lipnowski, E., Ravid, D., Shishkin, D., 2020. Persuasion via Weak Institutions. Working Paper.
Myerson, R.B., 1991. Game Theory - Analysis of Conflict. Harvard University Press.
Pai, M.M., Vohra, R., 2014. Optimal auctions with financially constrained buyers. J. Econom. Theory 150 (C), 383–425.
Queyranne, M., Schulz, A.S., 1994. Polyhedral Approaches to Machine Scheduling. Technical Report.
Salamanca, A., 2021. The value of mediated communication. J. Econom. Theory 192, 105191.
Shaked, M., Shanthikumar, J., 2007. Stochastic Orders. In: Springer Series in Statistics, Springer New York.
Vaidya, P., 1989. Speeding-up linear programming using fast matrix multiplication. In: 30th Annual Symposium on Foundations of Computer Science. pp. 332–337.