The Ronald O. Perelman Center for Political Science and Economics (PCPSE)
133 South 36th Street
Philadelphia, PA 19104-6297

pier@econ.upenn.edu
http://economics.sas.upenn.edu/pier

PIER Working Paper
24-025

Persuasion, Posteriors & Polymatroids

THÀNH NGUYEN
Purdue University

RAKESH VOHRA
University of Pennsylvania

August 1, 2024

Persuasion, Posteriors & Polymatroids

Thành Nguyen* Rakesh Vohra†

August 1, 2024

Abstract

Any mapping from states to signals induces a distribution over possible posteriors. For decision-making purposes, one is interested in parameters of those posteriors, such as their quartiles or whether they first-order stochastically dominate some given distribution. We show that a generalization of Gale's demand theorem can be used to characterize which distributions over possible posteriors with the requisite properties can be generated by some mapping from states to signals.

Keywords: Bayesian Persuasion, Information Design, Polymatroids

1 Introduction

The problem of Bayesian persuasion (see Kamenica and Gentzkow (2011) and Bergemann and Morris (2016)) is concerned with the properties of the posterior distribution over states induced by a given signal structure. For analytical tractability the focus has been on one-dimensional summary statistics such as the mean (see Dworczak and Martini (2019) and Vohra et al. (2023)) or the median of the resulting posterior (see Benoît and Dubra (2011)) or more generally the q^{th} quantile for any $q \in (0, 1)$ (see Kolotilin and Wolitzky (2024) and Yang and Zentefis (2022)).

The setting involves a finite state space, a prior distribution π over it, and a stochastic mapping of the states to signals. Each realized signal induces a posterior distribution over the states. For the moment, suppose one is interested in the q^{th} quantile of that posterior.

*Daniels School of Business, Purdue University, 403 W. State Street, West Lafayette, Indiana, 47906, United States. E-mail: nguye161@purdue.edu.

†Department of Economics & Department of Electrical and Systems Engineering University of Pennsylvania. Email: rvohra@seas.upenn.edu.

Different realized signals induce different posteriors, generating different q^{th} quantiles. Now, fix some distribution over q^{th} quantiles. Is there a mapping of states to signals that will induce this distribution over q^{th} quantiles? More generally, can one characterize the set of distributions of posterior quantiles that can be generated by some mapping of states to signals? Call such distributions of q^{th} quantiles **implementable**.

Benoît and Dubra (2011) characterize implementable distributions for the median under a uniform prior and finite state and signal space. Kolotilin and Wolitzky (2024) and Yang and Zentefis (2022) characterize it in the continuum case, for *any* prior distribution and *any* quantile. The proofs in Benoît and Dubra (2011) and Yang and Zentefis (2022) are lengthy. Kolotilin and Wolitzky (2020) provide a direct proof by showing that all implementable distributions of the posterior q^{th} quantiles are implemented by a single mapping from states to signals. We show that these results, for the finite case, are a straightforward consequence of Gale's demand theorem (Gale (1957)), which characterizes the existence of a feasible flow in a bipartite network. It is a modest generalization of Hall's matching Theorem (Hall (1935)). It has been employed in auction theory (Che et al. (2013)) and in Bayesian persuasion (Vohra et al. (2023)). Some well-known results in probability theory, such as Strassen (1965) and Holley (1974), are also consequences.

Our methodology provides new results that go beyond merely characterizing which distributions over q^{th}-quantiles are implementable. For example, one may be interested in not just a *particular* quantile of the realized posterior but all three quartiles of the resulting posterior. More generally, suppose we are given some finite set of distributions and require that each realized posterior first order stochastically dominate some distribution in the family. Our general framework also applies to this setting and provides a characterization of what distributions over posteriors can be implemented that satisfy these properties. These results show the versatility and applicability of Gale's theorem and its extension. Roughly speaking, this tool can be deployed to deal with any parameters of a distribution that can be encoded using a collection of linear constraints that are 'nested' in the appropriate way.

The next section begins with a restricted setting designed to illustrate the technique as clearly as possible. The subsequent section characterizes the distribution of implementable q-quantiles. The concluding section outlines the generalization of Gale's Theorem and two applications.

2 Notation and Example

Let the state space be $\{1, \ldots, n\}$ and π the prior probability distribution over it. Suppose an experiment that maps states to signals and assume there are as many signals as states. Each realized signal induces a posterior over the state space. Denote the posterior distribution over the states given signal j by $p(\cdot|j)$. To determine whether a given probability distribution over the state space can be induced as a possible posterior, it is enough to check if there is a solution to the following flow problem. Introduce a supply node i for each state $i \in \Sigma$ with supply π_i. Call the set of supply nodes S. For each signal i, introduce a demand node and call the set of demand nodes D. Let $x(i, j)$ be the flow from supply node i to demand node j. In the language of experiments this is the joint probability of observing state i and signal j simultaneously. Given a conditional probability distribution $p(i|j)$ for each signal $j \in \{1, \ldots, n\}$, we need to determine if there is a flow x satisfying the following:

$$\sum_{j \in D} x(i, j) = \pi_i \ \forall i \in S \tag{1}$$

$$\frac{x(i, j)}{\sum_{k \in S} x(k, j)} = p(i|j) \ \forall i \in S, j \in D \tag{2}$$

$$x(i, j) \geq 0 \ \forall i \in S, j \in D \tag{3}$$

Constraints (1) and (2) follow from Bayesian updating.

We are interested in whether a particular probability distribution over quantiles can be implemented by some mapping of states to signals. If $F(\cdot)$ is a distribution over $\{1, \ldots, n\}$, then, the q-quantile of F is the smallest $r \in \{1, \ldots, n\}$ such that $F(r) \geq q$. To allow the main idea of the proof to shine through, we will, as a first step, adopt a stricter notion of q-quantile: the smallest number r such that $F(r) = q$. Under this stricter notion n can never be a q-quantile because $F(n) = 1$ and it is also possible that no q-quantile exists. In Section 3, we relax this definition.

For each $i \in \{1, \ldots, n\}$ and a fixed $q \in (0, 1)$, we want to know if we can achieve the following: the probability that r is the q-quantile of the realized posterior is ρ_r (note $\rho_n = 0$). If so, then, ρ is implementable.

We can formulate the problem of verifying whether ρ is implementable as follows:

$$\sum_{r \in D} x(i, r) = \pi_i \ \forall i \in S \tag{4}$$

$$\sum_{i=r+1}^{n} x(i, r) = (1 - q)\rho_r \ r \in \{1, \ldots, n - 1\} \tag{5}$$

$$\sum_{i=1}^{r} x(i, r) = q\rho_r \ r \in \{1, \ldots, n - 1\} \tag{6}$$

$$x(i,j) \geq 0 \tag{7}$$

Constraints (5) and (6) follow from our definition of a q-quantile. The constraint $\sum_{i=1}^{n} x(i,r) = \rho_r$ is absent because it is implied by the sum of (5) and (6).

Problem (4-7) is a transportation problem (Hitchcock (1941)). As before, introduce a supply node i for each state i with supply π_i. For each $r \in \{1, \ldots, n-1\}$ introduce two demand nodes denoted r^+ and r^- with demands $(1-q)\rho_r$ and $q\rho_r$ respectively. Call them plus and minus nodes. For each supply node i introduce an arc directed from node i to demand node r^+ if $r+1 \leq i$. For each supply node i, introduce an arc directed from node i to demand node r^- if $r \geq i$. Problem (4-7) can now be phrased as follows: does there exist a flow from the supply nodes to the demand nodes along the given arcs such that all demand is satisfied and all supplies exhausted?

Gale's demand Theorem (Gale (1957)) gives a necessary and sufficient condition for the feasibility of the system (4-7). For any subset K of demand nodes, let $N(K)$ denote the set of supply nodes with at least one arc directed into an element of K. Then, a feasible flow exists iff for every subset of demand nodes K, the total supply in $N(K)$ is at least as large as the total demand in K. To state this formally, for any subset of demand nodes K, write K^+ for the plus nodes in K and K^- for the minus nodes. Then, the condition is

$$\sum_{i \in N(K)} \pi_i \geq \sum_{r \in K^+} (1-q)\rho_r + \sum_{r \in K^-} q\rho_r \ \forall K. \tag{8}$$

In principle, there are exponentially (in n) many of these conditions. However, it is easy to show that many will be redundant. Call a subset K of demand nodes **consecutive** if K is of the form $\{b^+, (b+1)^+, \ldots, (n-1)^+\}$ or $\{1^-, \ldots, a^-\}$.

LEMMA 2.1 *If constraint (8) holds for all consecutive subsets K of demand nodes, then it holds for all subsets K.*

Proof. To deny the conclusion is to admit the existence of a non-consecutive set K of demand nodes such that

$$\sum_{i \in N(K)} \pi_i - \sum_{r \in K^+} (1-q)\rho_r - \sum_{r \in K^-} q\rho_r < 0.$$

Call this difference the deficit. We show that if any subset of demand nodes has a negative deficit, there must be a consecutive subset of demand nodes with a negative deficit.

If K contains the plus nodes j^+ and $(j+r)^+$ for some r but not a^+, such that $j < a < j+r$, then including a^+ in K would leave the left-hand side of (8) unchanged but increase the right-hand side, thus producing another set with a negative deficit. Similarly, if there is a k^+ whose index exceeds the index of the largest plus node in K, including it would decrease the deficit. The same argument applies to the minus nodes. Therefore, we may assume that there exist indices a and b with

$$K = \{1^-, \dots, a^-\} \cup \{b^+, \dots, (n-1)^+\}.$$

If $a \geq b$, then $N(K)$ consists of all supply nodes and (8) holds. So, suppose that $a < b$. Then,

$$N(K) = \{1, \dots, a\} \cup \{b+1, \dots, n\} = N(K^-) \cup N(K^+).$$

Therefore, (8) must fail for at least one of K^- or K^+, which implies at least one consecutive subset with a negative deficit, a contradiction. ■

Given the Lemma it suffices to enforce (8) only for consecutive sets of demand nodes. When $K = \{r^+, \dots, (n-1)^+\}$ this would imply that

$$\sum_{j=r}^{n-1} \rho_j \leq \frac{\sum_{j=r+1}^{n} \pi_j}{1-q} \Rightarrow 1 - \sum_{j=1}^{r-1} \rho_j \leq \frac{\sum_{j=r+1}^{n} \pi_j}{1-q} \Rightarrow \sum_{j=1}^{r-1} \rho_j \geq \frac{\sum_{j=1}^{r} \pi_j - q}{1-q}.$$

When $K = \{1^-, \dots, r^-\}$ we conclude that

$$\sum_{j=1}^{r} \rho_j \leq \frac{\sum_{j=1}^{r} \pi_j}{q} \ \forall r.$$

This yields the following:

Theorem 2.1 *A distribution $\{\rho_r\}_{r=1}^{n}$ over q-quantiles is implementable iff*

$$\max\{\frac{\sum_{j=1}^{r+1} \pi_j - q}{1-q}, 0\} \leq \sum_{j=1}^{r} \rho_j \leq \min\{1, \frac{\sum_{j=1}^{r} \pi_j}{q}\} \ \forall r \leq n-1.$$

Theorem 2.1 tells us which distributions over quantiles (under the strict notion) are implementable but not the mapping from states to signals that will induce it. However, an extreme point solution to system (4-7) (when it exists) is easily found using the Northwest corner rule, and this rule characterizes the set of extreme points (see Section 8.1 of Brualdi (2006)).

2.1 Optimizing over ρ

The problem of maximizing a linear function of the ρ_js satisfying the constraints in Theorem 2.1 is an instance of maximizing a linear function over a generalized polymatroid. A generalized polymatroid is a polyhedron associated with a pair of real-valued functions $g(\cdot)$ and $f(\cdot)$ defined over subsets of $\{1, \ldots, n\}$ that are called paramodular:

1. $g(\cdot)$ is non-decreasing and supermodular,
2. $f(\cdot)$ is non-decreasing and submodular,
3. $g(\emptyset) = f(\emptyset) = 0$,
4. $f(S) - g(T) \geq f(S \setminus T) - g(T \setminus S)$ holds for all $S, T \subseteq \{1, \ldots, n\}$.

Given a paramodular pair (g, f), a generalized polymatroid is

$$\{y \in \Re^n : g(S) \leq \sum_{j \in S} y_j \leq f(S) \ \forall S \subseteq \{1, \ldots, n\}\}.$$

Optimizing a linear function over a generalized polymatroid can be accomplished using a greedy algorithm; see Chapter 14.5 in Frank (2011) for details. This also gives a characterization of the extreme points of a generalized polymatroid.

We now explain why the constraints in Theorem 2.1 correspond to a generalized polymatroid. Let $\mathcal{L}$ be a laminar family subsets of $\{1, \ldots, n\}$.[1] To eliminate some inessential possibilities, we assume that the maximal (in the sense of set inclusion) elements of $\mathcal{L}$ partition $\{1, \ldots, n\}$.

For each $S \in \mathcal{L}$ there are two non-negative integers $k(S)$ and $b(S)$ so that the following is feasible:

$$b(S) \leq \sum_{i \in S} y_i \leq k(S) \ \forall S \in \mathcal{L} \tag{9}$$

$$y_i \geq 0 \ \forall i \in \{1, \ldots, n\} \tag{10}$$

That the maximal elements of $\mathcal{L}$ partition $\{1, \ldots, n\}$ ensures that system (9-10) is bounded.

The problem of maximizing a linear function over (9-10) can be reduced to the problem of optimization over a generalized polymatroid for a suitable pair of paramodular functions $\{g(\cdot), f(\cdot)\}$. For each $S \subseteq \{1, \ldots, n\}$ set

$$f(S) = \max \sum_{i \in S} y_i$$

$$\text{s.t.} \ \sum_{i \in A} y_i \leq k(A) \ \forall A \in \mathcal{L}$$

[1] A family of subsets is laminar if, for any two members in the family, either they are disjoint or one is contained in the other. Budish et al. (2013) call the family hierarchical.

$$y_i \geq 0 \; \forall i \in \{1, \dots, n\}$$

and

$$g(S) = \min \sum_{i \in S} y_i$$

$$\text{s.t. } \sum_{i \in A} y_i \geq b(A) \; \forall A \in \mathcal{L}$$

$$y_i \geq 0 \; \forall i \in \{1, \dots, n\}$$

2.2 Monotone Function Intervals

Here, we remark that a problem closely related to characterizing implementable quantiles is also related to generalized polymatroids. It is the problem of characterizing the set of monotone functions that lie pointwise between two fixed monotone functions. It was studied in Yang and Zentefis (2024) for the continuum case. We focus of the 'discrete' case. The continuum case follows as the discretization becomes finer and finer via Helly's selection Theorem (Jenssen (2024)).

Let $\ell(\cdot)$ and $u(\cdot)$ be non-negative real valued non-decreasing functions defined on $\{0, 1, \dots, n\}$ such that $\ell(k) \leq u(k)$ for all $k \in \{0, 1, \dots, n\}$. The restriction of $\ell(\cdot)$, $u(\cdot)$ to being non-negative is for convenience only. Let $K(\ell, u)$ denote the set of monotone functions that lie between $\ell(\cdot)$ and $u(\cdot)$, i.e.,

$$K(\ell, u) = \{h : \{0, 1 \dots, n\} \to \Re : \ell(k) \leq h(k) \leq u(k) \; \forall k\}.$$

Now, any non-negative, non-decreasing function $h(\cdot)$ defined on $\{0, 1, \dots, n\}$ can be expressed $h(k) = \sum_{j=0}^{k} x_j$ where $x_j \geq 0$ for all $j = 0, 1, \dots, n$. Therefore, an equivalent description of $K(\ell, u)$ is the set of solutions to

$$\ell(k) \leq \sum_{j=1}^{k} x_j \leq u(k) \; \forall k = 0, 1, \dots, n$$

$$x_j \geq 0 \; j = 0, 1, \dots, n$$

Notice that the set of constraints forms a laminar family. Hence, the set of monotone functions that lie pointwise between two fixed monotone functions is a generalized polymatroid.

3 Relaxing The Notion of Quantile

We relax the strict notion of quantile used earlier and show that a variation of Theorem 2.1 still holds. Recall, that a q-quantile of a distribution $F(\cdot)$ over $\{1,\ldots,n\}$ is the smallest integer r such that $F(r) \geq q$. A necessary condition for a distribution $F(\cdot)$ to have a q-quantile of r is the existence of an $\epsilon > 0$ such that $F(r) \geq q$ and $F(r-1) \leq q-\epsilon$. We use this observation to characterize which distributions $\{\rho_r\}_{r=1}^{n}$ over q-quantiles are implementable.

Given an $\epsilon > 0$, a distribution $\{\rho_r\}_{r=1}^{n}$ over q-quantiles is implementable if there is a feasible solution to the following:

$$\sum_{j \in D} x(i,j) = \pi_i \ \forall i \in \{1,\ldots,n\} \tag{11}$$

$$\sum_{i=1}^{n} x(i,r) = \rho_r \ r \in \{1,\ldots,n\} \tag{12}$$

$$\sum_{i=1}^{r} x(i,r) \geq q\rho_r \tag{13}$$

$$\sum_{i=1}^{r-1} x(i,r) \leq (q-\epsilon)\rho_r \ r \in \{2,\ldots,n\} \tag{14}$$

$$x(i,r) \geq 0 \ \forall i,r \tag{15}$$

THEOREM 3.1 *System (11- 15) is feasible iff*

$$\max\{\frac{\sum_{i=1}^{k}\pi_i - q + \epsilon}{1-q+\epsilon}, 0\} \leq \sum_{r=1}^{k}\rho_r \leq \min\{\frac{1, \sum_{i=1}^{k}\pi_i}{q}\}, \ \forall r \leq n-1$$

The proof is an immediate consequence of a generalization of Gale's Theorem, which is discussed in the next section.

4 Generalizations

We consider two generalizations of the problem of determining which distributions over quantiles are implementable. In each case, the relevant characterization is easily obtained using a generalization of Gale's Theorem.

In the first generalization, for economy of exposition only, focus on the case of quartiles. Let D denote a set of possible quartiles. An element of D is a vector (r_1, r_2, r_3) with $r_1 < r_2 < r_3$. Here r_1 is the 25^{th} percentile, r_2 is the 50^{th} percentile while r_3 is the 75^{th}

percentile. Denote by $\rho(r_1, r_2, r_3)$ the 'desired' probability of realizing a posterior whose quartiles take values (r_1, r_2, r_3). Given D and a probability vector ρ over the elements of D, is there a mapping of states to signals that will induce the distribution ρ over D? This can be restated as whether the system below is feasible.

$$\sum_{j \in D} x(i,j) = \pi_i \;\forall i \in \{1, \ldots, n\} \tag{16}$$

$$\sum_{i=1}^{n} x(i,j) = \rho_j \;\forall j \in D \tag{17}$$

$$\sum_{i \leq r_1^j} x(i,j) \geq 0.25\rho_j \tag{18}$$

$$\sum_{i \leq r_2^j} x(i,j) \geq 0.5\rho_j \tag{19}$$

$$\sum_{i \leq r_3^j} x(i,j) \geq 0.75\rho_j \tag{20}$$

$$\sum_{i \leq r_1^j - 1} x(i,j) \leq (0.25 - \epsilon)\rho_j \tag{21}$$

$$\sum_{i \leq r_2^j - 1} x(i,j) \leq (0.5 - \epsilon)\rho_j \tag{22}$$

$$\sum_{i \leq r_3^j - 1} x(i,j) \leq (0.75 - \epsilon)\rho_j \tag{23}$$

$$x(i,j) \geq 0 \;\forall i \in \{1, \ldots, n\},\; j \in D \tag{24}$$

THEOREM 4.1 *Let D be a finite set of possible quartile vectors r and ρ a probability distribution over D. There is a mapping of states to signals such that the probability of the realized posterior having quartiles $r \in D$ is ρ_r iff,*

$$0.25 \sum_{j: r_1^j \leq k \leq r_2^j - 1} \rho_j + 0.5 \sum_{j: r_2^j \leq k \leq r_3^j - 1} \rho_j + 0.75 \sum_{j: r_3^j \leq k \leq n-1} \rho_j \leq \sum_{i=1}^{k} \pi_i \;\forall k \leq n-1$$

$$(0.25-\epsilon) \sum_{j: k \leq r_1^j - 1} \rho_j + (0.5-\epsilon) \sum_{j: r_1^j \leq k \leq r_2^j - 1} \rho_j + (0.75-\epsilon) \sum_{j: r_2^j \leq k \leq r_3^j - 1} \rho_j + \sum_{j: r_3^j \leq k} \rho_j \geq \sum_{i=1}^{k} \pi_i \;\forall k \leq n-1$$

Now, consider a different yet related question. First, we recall the notion of first order stochastic dominance. A distribution $H(\cdot)$ over $\{1, \ldots, n\}$ is said to first-order stochastically dominate (FOSD) distribution $W(\cdot)$ if

$$H(r) \leq W(r) \; \forall r \leq n-1.$$

If the corresponding density functions are denoted $h(\cdot)$ and $w(\cdot)$ respectively, the FOSD condition can be written as:

$$\sum_{j=1}^{r} h(j) \leq \sum_{j=1}^{r} w(j) \; \forall r \leq n-1$$

$$\sum_{j=1}^{n} h(j) = 1$$

If H FOSD W, then each q-quantile of H is no larger than the corresponding q-quantile of W.

Now, let D be some finite collection of probability distributions. Fix a probability distribution $\rho(\cdot)$ over the elements of D. Thus, $\rho(W)$ is the probability assigned to distribution W in D. Call $\rho(\cdot)$ implementable if there is a mapping of states to signals such that the probability of the realized posterior FOSD distribution $W \in D$ is $\rho(W)$.

THEOREM 4.2 *Let D be a finite set of probability distributions and ρ a probability distribution over D. Then, $\rho(\cdot)$ is implementable iff*

$$\sum_{W \in D} \rho(W) W(r) \geq \sum_{i=1}^{r} \pi_i \; \forall r = 1, \ldots, n.$$

Theorems 3.1, 4.1, and 4.2 are corollaries of a generalization of Gale's Theorem (see Theorem 6 in Che et al. (2013)) described below. The proof of Theorem 4.1 appears in the appendix. The proofs of Theorem's 3.1 and 4.2 are similar and are omitted.

4.1 Generalization of Gale's Theorem

We provide a formal statement of the generalization of Gale's Theorem that is independent of the auction context surrounding it in Che et al. (2013), which we hope will encourage its wider use. For completeness, we provide a proof that differs from the one in (Che et al. (2013)). Its virtue is its brevity.

Let $\{1, \ldots, n\}$ be the set of supply nodes, each with a supply of π_i. Let D be the set of demand nodes, and associated with each demand node $j \in D$ is a pair of paramodular functions $\{g_j(\cdot), f_j(\cdot)\}$ defined on subsets of $\{1, \ldots, n\}$.

The associated flow problem is:

$$\sum_{i=1}^{n} x_{ij} = \pi_i \ \forall i \in \{1,\dots,n\} \tag{25}$$

$$g_j(S) \le \sum_{i\in S} x_{ij} \le f_j(S) \ \forall S \subseteq \{1,\dots,n\} \ \forall j \in D \tag{26}$$

$$x_{ij} \ge 0 \ \forall i \in \{1,\dots,n\}, j \in D \tag{27}$$

The following is Theorem 6 in (Che et al. (2013)) stripped of its auction context.

THEOREM 4.3 *System (25-27) is feasible iff. for all disjoint* $T^+, T^- \subseteq \{1,\dots,n\}$

$$\sum_{j\in D} f_j(T^-) - \sum_{i\in T^-} \pi_i \ge 0 \ge \sum_{j\in D} g_j(T^+) - \sum_{i\in T^+} \pi_i.$$

Proof. If the system (25-27) is infeasible, by the Farkas lemma, there exists a feasible solution to its alternative:

$$\sum_{i=1}^{n} \pi_i s_i + \sum_{j\in D} \sum_{S\subseteq\{1,\dots,n\}} f_j(S) y_j(S) - \sum_{j\in D} \sum_{S\subseteq\{1,\dots,n\}} g_j(S) z_j(S) < 0 \tag{28}$$

$$s_i + \sum_{S\ni i} y_j(S) - \sum_{S\ni i} z_j(S) \ge 0 \ \forall i = 1,\dots,n \ \forall j \in D \tag{29}$$

$$y_j(S), z_j(S) \ge 0 \ \forall S \subseteq \{1,\dots,n\} \ \forall j \in D \tag{30}$$

If (s^*, y^*, z^*) is a feasible solution to (28-30), by scaling we can assume

$$-1 \le s_i \le 1 \ \forall i \in \{1,\dots,n\}. \tag{31}$$

$$y_j(S), z_j(S) \le 1 \ \forall j \in D \ \forall S \subseteq \{1,\dots,n\}. \tag{32}$$

Therefore, we can assume that (s^*, y^*, z^*) is an extreme point of $(29-32)$. We show that the extreme points of $(29-32)$ are integral.

A standard uncrossing argument implies that $\{S : y_j^*(S) > 0\}$ and $\{S : z_j^*(S) > 0\}$ are laminar families. Now, the incidence matrix of the union of two laminar families is totally unimodular. If we append to a totally unimodular matrix an identity matrix, constraints (31-32), the augmented matrix is also totally unimodular. Hence, the basis matrix associated with the extreme point is totally unimodular, implying that (s^*, y^*, z^*) is integral. In fact, each $s_i \in \{-1, 0, +1\}$ for all $i \in \{1,\dots,n\}$.

Let $T^+ = \{i \in \{1,\dots,n\} : s_i^* = 1\}$ and $T^- = \{i \in \{1,\dots,n\} : s_i^* = -1\}$. We now determine the values of (y^*, z^*) by selecting them to minimize $\sum_{j\in D}\sum_{S\subseteq\{1,\ldots,n\}} f_j(S)y_j(S) - \sum_{j\in D}\sum_{S\subseteq\{1,\ldots,n\}} g_j(S)z_j(S)$ subject to (29-32) holding s^* fixed. Stating this problem explicitly:

$$\min \sum_{j\in D}\sum_{S\subseteq\{1,\ldots,n\}} f_j(S)y_j(S) - \sum_{j\in D}\sum_{S\subseteq\{1,\ldots,n\}} g_j(S)z_j(S)$$

$$\text{s.t. } \sum_{S\ni i} y_j(S) - \sum_{S\ni i} z_j(S) \geq -1 \ \forall i \in T^+ \ \forall j \in D$$

$$\sum_{S\ni i} y_j(S) - \sum_{S\ni i} z_j(S) \geq 1 \ \forall i \in T^- \ \forall j \in D$$

$$\sum_{S\ni i} y_j(S) - \sum_{S\ni i} z_j(S) \geq 0 \ \forall i \notin T^+ \cup T^- \ \forall j \in D$$

$$y_j(S), z_j(S) \geq 0 \ \forall S \subseteq \{1,\ldots,n\} \ \forall j \in D$$

We will conclude that $y^*_j(T^-) = 1$ for all $j \in D$ and $y^*_j(S) = 0$ whenever $S \neq T^-$. Similarly, $z^*_j(T^+) = 1$ for all $j \in D$ and $z^*_j(S) = 0$ whenever $S \neq T^+$. Recall, we can assume that $\{S : y^*_j(S) > 0\}$ and $\{S : z^*_j(S) > 0\}$ are laminar and $y^*_j(S), z^*_j(S) \in \{0,1\}$ for all j and S.

1. For all $j \in D$ and $S \subseteq \{1,\ldots,n\}$ we may assume that $y^*_j(S)z^*_j(S) = 0$. If not, decrease $y^*_j(S)$ and $z^*_j(S)$ by 1. Feasibility is preserved and objective function value declines by $f_j(S) - g_j(S) \geq 0$.

2. Subadditivity of $f_j(\cdot)$ implies we can eliminate the possibility of $A, B \in \{S : y^*_j(S) > 0\}$ such that $A \cap B = \emptyset$. Therefore, feasibility requires that $y^*_j(S) = 1$ for some S containing T^-.

3. Given that $f_j(\cdot)$ is non-decreasing, $y^*_j(S) = 0$ whenever S is a strict superset of T^-. Hence, $y^*_j(T^-) = 1$.

4. Given $y^*_j(T^-) = 1$, we can set $y^*_j(S) = 0$ for all $S \subset T^-$ without violating feasibility.

A similar argument applies to the $z_j(S)$ variables. Hence, infeasibility implies there exist disjoint sets $T^+, T^- \subseteq \{1,\ldots,n\}$ such that

$$\sum_{i\in T^+} \pi_i - \sum_{i\in T^-} \pi_i + \sum_{j\in D} f_j(T^-) - \sum_{j\in D} g_j(T^+) < 0.$$

■

We consider Theorem 4.3 to be folklore. Special cases of it, such as the matroid partition

Theorem, Edmonds (1965), are known. The proof given here is a variation of a proof of a closely related result.

We state a variant of Theorem 4.3 (also folklore) that is useful. Associate with each $j \in D$ a laminar family of subsets of $\{1, \ldots, n\}$ denoted $\mathcal{L}^j$. For each $S \in \mathcal{L}^j$ there are two non-negative integers $k^j(S)$ and $b^j(S)$ so that the following is feasible:

$$b^j(S) \leq \sum_{i \in S} y_i \leq k^j(S) \ \forall S \in \mathcal{L}^j$$

$$y_i \geq 0 \ \forall i \in \{1, \ldots, n\}$$

The system whose feasibility will be of interest is

$$\sum_{i=1}^{n} x_{ij} = \pi_i \ \forall i \in \{1, \ldots, n\} \tag{33}$$

$$b^j(S) \leq \sum_{i \in S} x_{ij} \leq k^j(S) \ \forall S \in \mathcal{L}^j \ \forall j \in D \tag{34}$$

$$x_{ij} \geq 0 \ \forall i \in \{1, \ldots, n\}, j \in D \tag{35}$$

Verifying feasibility of (33-35) can be reduced to verifying the feasibility of (25-27) for a suitable pair of paramodular functions $\{g_j(\cdot), f_j(\cdot)\}_{j \in D}$. For each $j \in D$ set

$$f_j(S) = \max \sum_{i \in S} y_i$$

$$\text{s.t.} \ \sum_{i \in A} y_i \leq k^j(A) \ \forall A \in \mathcal{L}^j$$

$$y_i \geq 0 \ \forall i \in \{1, \ldots, n\}$$

and

$$g_j(S) = \min \sum_{i \in S} y_i$$

$$\text{s.t.} \ \sum_{i \in A} y_i \geq b^j(A) \ \forall A \in \mathcal{L}^j$$

$$y_i \geq 0 \ \forall i \in \{1, \ldots, n\}.$$

References

Benoît, J.-P. and J. Dubra (2011): "Apparent Overconfidence," *Econometrica*, 79, 1591–1625.

Bergemann, D. and S. Morris (2016): "Information Design, Bayesian Persuasion, and Bayes Correlated Equilibrium," *American Economic Review*, 106, 586–91.

Brualdi, R. A. (2006): *Combinatorial Matrix Classes*, Encyclopedia of Mathematics and its Applications, Cambridge University Press.

Budish, E., Y.-K. Che, F. Kojima, and P. Milgrom (2013): "Designing Random Allocation Mechanisms: Theory and Applications," *American Economic Review*, 103, 585–623.

Che, Y.-K., J. Kim, and K. Mierendorff (2013): "Generalized Reduced-Form Auctions: A Network-Flow Approach," *Econometrica*, 81, 2487–2520.

Dworczak, P. and G. Martini (2019): "The Simple Economics of Optimal Persuasion," *Journal of Political Economy*, 127, 1993–2048.

Edmonds, J. (1965): "Minimum Partition of a Matroid into Independents Subsets," *Journal of Research of the National Bureau of Standards*, 69, 67–72.

Frank, A. (2011): *Connections in Combinatorial Optimization*, Oxford Lecture Series in Mathematics and Its Applications, OUP Oxford.

Gale, D. (1957): "A theorem on flows in networks." *Pacific Journal of Mathematics*, 7, 1073–1082.

Hall, P. (1935): "On Representatives of Subsets," *Journal of the London Mathematical Society*, s1-10, 26–30.

Hitchcock, F. L. (1941): "The Distribution of a Product from Several Sources to Numerous Localities," *Journal of Mathematics and Physics (MIT)*, 20, 224–230.

Holley, R. (1974): "Remarks on the FKG inequalities," *Communications in Mathematical Physics*, 36, 227–231.

Jenssen, H. K. (2024): "A selection theorem for essentially regulated functions of several variables," *Journal of Mathematical Analysis and Applications*, 534.

Kamenica, E. and M. Gentzkow (2011): "Bayesian Persuasion," *American Economic Review*, 101, 2590–2615.

Kolotilin, A. and A. Wolitzky (2020): "Assortative Information Disclosure," Discussion Papers 2020-08, School of Economics, The University of New South Wales.

——— (2024): "Distributions of Posterior Quantiles via Matching," .

Strassen, V. (1965): "The Existence of Probability Measures with Given Marginals," *The Annals of Mathematical Statistics*, 36, 423 – 439.

Vohra, A., J. Toikka, and R. Vohra (2023): "Bayesian persuasion: Reduced form approach," *Journal of Mathematical Economics*, 107, 102863.

Yang, K. H. and A. Zentefis (2022): "Distributions of Posterior Quantiles and Economic Applications," *Working Paper.*

Yang, K. H. and A. K. Zentefis (2024): "Monotone Function Intervals: Theory and Applications," *American Economic Review*, 114, 2239–70.

Appendix: Proof of Theorem 4.1

In observing the constraints (16-24), we see that we can associate with each $j \in D$ two laminar families. The first is:

$$\sum_{i=1}^{n} x(i,j) \geq \rho_j \ \forall j \in D$$

$$\sum_{i \leq r_1^j} x(i,j) \geq 0.25\rho_j$$

$$\sum_{i \leq r_2^j} x(i,j) \geq 0.5\rho_j$$

$$\sum_{i \leq r_3^j} x(i,j) \geq 0.75\rho_j$$

This gives rise to the following supermodular function for each $j \in D$:

$$g_j(S) = \min \sum_{i \in S} y_i$$

$$\text{s.t.} \ \sum_{i \leq r_1^j} y_i \geq 0.25\rho_j$$

$$\sum_{i \le r_2^j} y_i \ge 0.5\rho_j$$

$$\sum_{i \le r_3^j} y_i \ge 0.75\rho_j$$

$$\sum_{i=1}^{n} y_i \ge \rho_j$$

$$y_i \ge 0 \ \forall i$$

A straightforward computation shows that:

1. $g_j(\{1, \ldots, k\}) = 0.25\rho_j$ if $r_1^j \le k \le r_2^j - 1$
2. $g_j(\{1, \ldots, k\}) = 0.5\rho_j$ if $r_2^j \le k \le r_3^j - 1$
3. $g_j(\{1, \ldots, k\}) = 0.75\rho_j$ if $r_3^j \le k \le n - 1$
4. $g_j(\{1, \ldots, k\}) = \rho_j$ if $k = n$.
5. $g_j(S) = 0$ in all other cases.

Similarly, for each $j \in D$ we have the following submodular function:

$$f_j(S) = \max \sum_{i \in S} y_i$$

$$\text{s.t.} \quad \sum_{i \le r_1^j - 1} y_i \le (0.25 - \epsilon)\rho_j$$

$$\sum_{i \le r_2^j - 1} y_i \le (0.5 - \epsilon)\rho_j$$

$$\sum_{i \le r_3^j - 1} y_i \le (0.75 - \epsilon)\rho_j$$

$$\sum_{i=1}^{n} y_i \le \rho_j$$

$$y_i \ge 0 \ \forall i$$

A straightforward computation shows that

1. $f_j(S) = (0.25 - \epsilon)\rho_j$ if $\max_{i \in S} i \le r_1^j - 1$
2. $f_j(S) = (0.5 - \epsilon)\rho_j$ if $r_1^j \le \max_{i \in S} i \le r_2^j - 1$

3. $f_j(S) = (0.75 - \epsilon)\rho_j$ if $r_2^j \leq \max_{i \in S} i \leq r_3^j - 1$

4. $f_j(S) = \rho_j$ if $r_3^j \leq \max_{i \in S} i$

From Theorem 4.3, we know that to ensure feasibility:

$$\sum_{j \in D} f_j(T^-) - \sum_{i \in T^-} \pi_i \geq 0 \geq \sum_{j \in D} g_j(T^+) - \sum_{i \in T^+} \pi_i \tag{36}$$

for all disjoint $T^-, T^+ \subseteq \{1, \ldots, n\}$.

Focus on the right-hand side of (36) first. From the properties of $g_j(\cdot)$, it's clear that the only sets T^+ that yield a non-trivial condition are of the form $\{1, \ldots, k\}$ for $k \leq n-1$. This yields:

$$0.25 \sum_{j: r_1^j \leq k \leq r_2^j - 1} \rho_j + 0.5 \sum_{j: r_2^j \leq k \leq r_3^j - 1} \rho_j + 0.75 \sum_{j: r_3^j \leq k \leq n-1} \rho_j - \sum_{i=1}^{k} \pi_i \leq 0.$$

Now, consider the left hand side of (36). Suppose there is a $t \notin T^{-+}$ such that $t < \max_{i \in T^-} i$. If we replace T^- by $T^- \cup \{t\}$ in the left-hand side of (36), this can only lower it. Hence, we can restrict ourselves to sets T^- such that $T^- = \{1, \ldots, k\}$ or $T^- = \{k, \ldots, n\}$. The first kind yields:

$$(0.25 - \epsilon) \sum_{j: k \leq r_1^j - 1} \rho_j + (0.5 - \epsilon) \sum_{j: r_1^j \leq k \leq r_2^j - 1} \rho_j + (0.75 - \epsilon) \sum_{j: r_2^j \leq k \leq r_3^j - 1} \rho_j + \sum_{j: r_3^j \leq k} \rho_j - \sum_{i=1}^{k} \pi_i \geq 0.$$

If we set $T^- = \{k+1, \ldots, n\}$ and $T^+ = \{1, \ldots, k\}$, this yields:

$$\sum_{j \in D} \rho_j - \sum_{i=k+1}^{n} \pi_i \geq 0.25 \sum_{j: r_1^j \leq k \leq r_2^j - 1} \rho_j + 0.5 \sum_{j: r_2^j \leq k \leq r_3^j - 1} \rho_j + 0.75 \sum_{j: r_3^j \leq k \leq n-1} \rho_j - \sum_{i=1}^{k} \pi_i$$

$$\Rightarrow 2 \sum_{i=1}^{k} \pi_i \geq 0.25 \sum_{j: r_1^j \leq k \leq r_2^j - 1} \rho_j + 0.5 \sum_{j: r_2^j \leq k \leq r_3^j - 1} \rho_j + 0.75 \sum_{j: r_3^j \leq k \leq n-1} \rho_j$$

This is redundant given earlier inequality.