Selected Facts

Nemanja Antic[*] Archishman Chakraborty[†]

This version: September, 2022.

PRELIMINARY AND INCOMPLETE SEMINAR DRAFT.

Abstract

An expert wishes to persuade a receiver to accept or reject a proposal. She has relevant information about multiple aspects of the decision and must select a certain number of facts in order to make her case. We provide a general characterization of the expert's optimal reporting policy and identify necessary and sufficient conditions under which she obtains her ideal decisions. We also characterize the expert's optimal strategy when she cannot commit to a reporting policy and faces incentive constraints. We obtain our results by casting the expert's problem as a matching problem on an appropriately chosen bipartite graph.

JEL Classification: C72, D71, D72, D82.

Keywords: persuasion, matching, evidence, cheap talk.

[*]Managerial Economics and Decision Sciences Department, Kellogg School of Management, Northwestern University, Evanston, Illinois; nemanja.antic@kellogg.northwestern.edu.

[†]Finance Department, Syms School of Business, Yeshiva University, New York, New York; archishman@yu.edu.

1 Introduction

In the Fall of 1945, reports of adverse and lingering health effects were emanating from Japan, attributed to the atomic bombs dropped on Hiroshima and Nagasaki. General Lesley Groves, head of the Manhattan Project, fought the allegations. Studies ordered by him using Geiger counters, and conducted at bomb test sites in the U.S. as well as at Hiroshima, had found negligible radioactive fallout at ground zero. These facts had a role in the U.S. denying the charges and it led to a decline in coverage of the story in the mainstream media. But these tests were not designed to uncover all the facts. Scientists and doctors involved with the Manhattan Project were aware of the health effects of radiation. The likely culprit was gamma ray radiation but detecting this needed more elaborate tests. Later historians contend that Gen. Groves himself knew this fact as early as 1943. But the Manhattan Project was organizationally so secretive and compartmentalized that even the U.S. administration was not aware of the radiation issues at the time the strikes were ordered by President Truman.[1]

In this paper, we consider the problem of an expert (the sender) who selects facts in order to persuade an uninformed observer (the receiver) to approve of a decision recommended (or taken) by the sender. The decision is to either accept or reject a proposal. The sender is privately informed about multiple aspects that are relevant for the decision. Relative to the receiver, she is biased in favor of accepting the proposal. She must (truthfully) reveal a minimum number of (indisputable) facts in her report that will subsequently be scrutinized by the receiver. Her design problem is to select the facts that will be revealed, and those that will be concealed, as a function of her private information, while making sure the receiver always approves of her decision and never wants to overturn it and take a different decision.

Such design problems seem pervasive. Politicians soliciting votes, or legislators arguing in parliament, select their facts as they make their case. Managers reporting to the board of directors, or the investing public, often submit evidence and projections together with recommendations. Because of their expertise about internal operations, they may have choice in what they present. For the same reason, organizations may be able to manage the information they disclose in order to meet regulatory requirements, and experts testifying in court can select the facts they choose to emphasize.

[1] Source: "The Black Reporter Who Exposed a Lie About the Atomic Bomb," New York Times, August 9, 2021 https://www.nytimes.com/2021/08/09/science/charles-loeb-atomic-bomb.html

We show that the sender must select her facts in a manner that pools states where she wants to accept the proposal but the receiver does not, with states where they both do. This is formally equivalent to a matching problem on a suitably chosen bipartite graph. One side of the graph corresponds to states where the sender and receiver agree that the proposal should be accepted whereas the other side corresponds to states where they have a conflict of interest. Each edge is an argument the sender can make, a set of facts she can reveal, that will persuade the receiver to accept her recommended decision. A matching on such a graph defines a reporting strategy for the sender.

We use Hall's marriage theorem (1935) to identify a sufficient condition under which the sender can persuade the receiver to approve the sender's own ideal decision in every state. This condition states that there must be enough diversity in the different ways the sender can select her facts in order to pool agreement states where both agree on the decision with states where there is conflict. When it is met, the sender subverts the receiver's attempt to monitor her and implements her own unconstrained optimal decision rule. There exists a perfect matching on the associated bipartite graph that can be used to construct the sender's ideal reporting policy. We provide examples of environments where it is met and the sender's ideal policy takes a natural and simple form.

Our sufficient condition is also necessary. When it is not met, the sender compromise and give up on her own ideal decision in some states. Such a constrained optimal reporting policy is equivalent to a maximal matching on a weighted bipartite graph. We identify the expected payoffs to both the sender and receiver from such a policy, for very general preferences and priors.

These results are derived under the assumption that the sender can commit to a reporting policy. She designs an experiment ex ante that reveals certain facts as a function of the state. Commitment is a standard assumption in the literature on persuasion and information design (Kamenica and Gentzkow, 2011) and it seems to fit many of the applications we have in mind. It is also a harmless assumption when a subversive reporting policy exists since the sender has no incentive to deviate from a strategy that results in her own ideal decision being taken in every state.

When subversion is not possible, the sender may have an incentive to depart from her policy after learning the state. She may want change her recommended decision while presenting some justifying facts. Using our graph-theoretic framework, we analyze the sequential equilibria of such a communication game. We provide necessary and sufficient conditions under which the sender-optimal equilibrium implements the receiver's ideal decisions in every state. Under some simplifying assumptions, we also fully characterize all equilibrium outcomes in an additive model where the sender always wants to accept the proposal while the receiver wants to do so if the average value of the aspects is large enough. The sender will not always elect to reveal the strongest fact in her favor in the sender-optimal equilibrium of this model. The receiver must consider it possible that the sender is understating her case.

The most closely related work to this paper is Glazer and Rubinstein (2004). They consider a situation where a sender communicates with a receiver following which the receiver selects one aspect to verify. Their focus is receiver-optimal mechanism design. We analyze the mirror-image case of sender-optimal information design. In our setup, the sender chooses an optimal rule of selecting facts in order to persuade a receiver and the latter does not have the ability to obtain independent facts on his own. Nonetheless, the two models are related and our additive model allows a close comparison that we provide later in the paper.

Fishman and Hagerty (1989) consider a situation where a sender in possession of multiple informative binary signals about a state can reveal only one. The sender has state independent preferences and she wants the receiver to think that she has many high signals. They show the receiver will benefit from limiting the sender's discretion about which signal to reveal. It may be optimal for the receiver to arrange the signals in an arbitrary order and infer from the fact that the kth signal is revealed that all signals with an index $k' < k$ were unfavorable to the sender. This emphasis on receiver-optimal equilibria is the opposite of ours. We focus on sender-optimal information design, or sequential equilibria of a communication game in which the receiver forms beliefs in a Bayes-consistent manner but otherwise does not have an active role.

Incomplete Literature Review.

2 Persuasion with selected facts

2.1 Model

A sender ("she") recommends a decision d to either accept a proposal ($d = a$) or reject it ($d = r$), as a function of her private information $x \in X$. She also prepares a report $m \in \mathcal{M}$. Her recommendation d and report m are scrutinized subsequently by an uninformed receiver ("he"). The receiver has a conflict of interest with the sender and may not agree with the sender's recommendation.

Let the receiver's payoff from accepting the proposal equal $u(x) > 0$ when $x \in A$, for some non-empty $A \subseteq X$; with $u(x) < 0$ otherwise. The sender's payoff from accepting the proposal is equal to $v(x) > 0$ for $x \in A \cup C \subseteq X$, $A \cap C = \emptyset$; with $v(x) < 0$ otherwise. The payoff from rejecting the proposal is normalized to zero for both the sender and the receiver. Let $R = X \setminus (A \cup C)$.

When $x \in R$ both the sender and the receiver prefer to reject the proposal, whereas when $x \in A$ both prefer to accept it. When $x \in C$ the sender prefers to accept the proposal whereas the receiver prefers to reject it. The set C describes the set of states where there is a conflict of interest between the sender and the receiver. The functions u and v provide the cardinal values of accepting the proposal to each player in each state.

Suppose $X = \times_{i=1,...,n} X_i$, $n \geq 1$. Each component x_i of the state of the world $x = (x_1, ..., x_n)$ is an *aspect* that is relevant for the decision. Each X_i is a finite set and p is a probability measure that represents the receiver's priors on X. Let $p(x)$ denote the probability of state x.

Given a report m from the sender and a recommendation $d \in \{a, r\}$, the receiver's preferences imply he (weakly) prefers to follow the sender's recommendation if and only if

$$\mathbb{E}[u \mid m, a] \geq 0 \geq \mathbb{E}[u \mid m, r]. \tag{1}$$

The sender must make a recommendation d and prepare her report m in a way that ensures (1) holds. This obedience constraint (Kamenica and Gentzkow, 2011) ensures the receiver has the incentive to follow the sender's recommendation.

As an alternative specification that is formally equivalent, we can allow the sender to take the decision herself while making sure the receiver has no incentive to overturn her decision. Under this interpretation, (1) captures a notion of deniability (Antic, Chakraborty and Harbaugh, 2022). When the sender has decision rights but must comply with regulations, or is subject to ex post public scrutiny, meeting this constraint allows her to maintain deniability that she may not have served the public interest. Doing so avoids interventions, lawsuits, protests or other penalties. In what follows, when we say the sender chooses a particular decision $d \in \{a, r\}$ we mean, interchangeably, either that she recommends that decision or that she takes it herself.

When $u(x) = +1$ for all $x \in A$ and $u(x) = -1$ otherwise, (1) becomes

$$\Pr[A \mid m, a] \geq \frac{1}{2} \geq \Pr[A \mid m, r]. \tag{2}$$

In this case the receiver wishes to avoid mistakes, i.e., choosing $d = r$ for $x \in A$ and $d = a$ for $x \notin A$. He cares equally about both kinds of mistakes. The corresponding deniability constraint (2) is similar to the "balance of probabilities" standard of proof used in for civil cases in the U.S.. When it is met, the balance of probabilities favors acquitting the sender. While we allow for more or less demanding standards for acquittal, such as probable cause or reasonable doubt, and allow the standard to be state dependent, this balance of probabilities special case will be important in what follows.

We impose the following restriction on the reports the sender is allowed to send. In each state $x = (x_1, ..., x_n)$, the report $m \in \mathcal{M}$ must (perfectly) reveal the values of at least k aspects, $k \in \{0, ..., n\}$. This defines a set of messages $\mathcal{M}(x) \subseteq \mathcal{M}$ that the sender can send in state x. Each message $m \in \mathcal{M}(x)$ must reveal the realized value of a k-dimensional vector of aspects x_K, for some $K \subseteq \{1, ..., n\}$ with $|K| = k$. The vector x_K describes a set of facts the sender is obligated to disclose truthfully. Let x_{-K} denote the remaining facts with $x = (x_K, x_{-K})$.

A reporting strategy σ for the sender picks a probability distribution $\sigma(x) \in \Delta(\mathcal{M}(x))$ for each $x \in X$. Let *supp* $\sigma(x)$ denote the support of $\sigma(x)$. For each report in *supp* $\sigma(x)$ the sender also chooses a decision $d \in \{a, r\}$. Let $supp_a\ \sigma(x)$ be the set of reports in *supp* $\sigma(x)$ which result in $d = a$, and similarly let $supp_r\ \sigma(x)$ denote the set of reports that result in $d = r$. Let $supp_a\ \sigma = \cup_{x \in X} supp_a\ \sigma(x)$ and similarly define $supp_r\ \sigma$. The sender commits to a reporting strategy σ in order to maximize her ex ante expected payoff, subject to decisions meeting the constraint (1).

This defines the sender's problem $\mathcal{P} = \langle \{A, C, R\}, \{p, u, v\}, n, k \rangle$ of selecting facts. The case $k = 0$ corresponds to a standard (though possibly multi-dimensional) persuasion environment. In comparison, when $k > 0$ we have a constrained persuasion problem because the sender faces an additional burden of selecting k facts that she must reveal in her report. She has the freedom to select these facts and to send any additional information that may reveal more. We suppose that each $\mathcal{M}(x)$ is rich enough so that information transmission is never constrained by the availability of messages.[2]

[2] Since the sender commits to a reporting strategy, it does not matter for our results if we allow her to reveal exactly k facts or to reveal more if she so wishes. We maintain the latter assumption throughout the paper because

Relative to the sender, the receiver has a passive role in our model. The sender's report m and recommendation (or decision) d can only be scrutinized by him ex post. He cannot request or obtain particular facts on his own, or design ex ante any other aspect of communication and evidence production. The mechanism design problem where the receiver designs a mechanism ex ante, that requests or verifies particular facts as a function of the sender's message, has been analyzed by Glazer and Rubinstein (2004). We discuss the connections to their work in Section 3, where we consider the case where the sender does not commit to a reporting strategy and instead engages in cheap talk on the facts that she does not reveal.

2.2 Subversive reporting strategies

Fix the sender's problem $\mathcal{P}$. Consider a bipartite graph $G = \{A \cup C, E\}$, with $A \cup C$ the set of vertices and E the set of edges, such that an edge $\{x, x'\} \in E$ that connects vertices $x \in C$ and $x' \in A$ must have $x_K = x'_K$ for some $K \subseteq \{1, ..., n\}$ with $|K| = k$. Thus, two vertices are connected by an edge if and only if (at least) k aspects of the vertices take the same values. For $S \subseteq A \cup C$, let $N(S) = \{x' \in A \cup C \mid \{x, x'\} \in E,\ x \in S\}$ denote the *neighbors* of S.

Call $M \subseteq E$ a *matching* if $\{x, x'\} \in M$ implies $\{x, x''\} \notin M$ and $\{x', x''\} \notin M$ for all $x'' \in A \cup C$. A vertex $x \in A \cup C$ is matched by M if there exists $x' \in A \cup C$ such that $\{x, x'\} \in M$; otherwise x is unmatched by M. A matching M on G is *C-perfect* if every $x \in C$ is matched by M. A matching M on G is *perfect* if every $x \in A \cup C$ is matched by M.

For our first result, we ask if it is possible for the sender to obtain her ideal payoffs, in the special case where priors are uniform and deniability corresponds to the balance of probabilities standard (2). To obtain her ideal payoffs, the sender must choose $d = a$ whenever $x \in A \cup C$ and $d = r$ otherwise. If there is a reporting strategy that makes it possible to implement this decision rule while satisfying the deniability constraint, then the sender avoids any burden of scrutiny by the receiver. She subverts the receiver's agenda and implements her own unconstrained optimal decision rule. We call such a reporting strategy a *subversive reporting strategy.*

Proposition 1 *Assume (i) all $x \in A \cup C$ are equiprobable and (ii) the deniability constraint is given by the balance of probabilities standard (2). A subversive reporting strategy exists if and only if*

$$|N(S)| \geq |S| \text{ for all } S \subseteq C. \tag{HC}$$

Proof.

it is convenient for the proofs. We discuss how this distinction may matter in Section 3 where we consider situations without commitment.

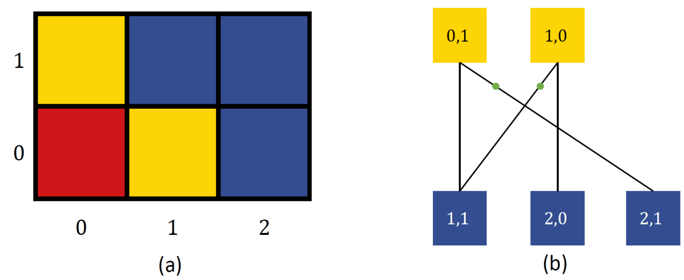

Figure 1: A perfect matching on G.

Step 1 (‘if’). By Hall’s theorem (1935) a C-perfect matching exists on G if and only if (HC) holds. Suppose (HC) holds and let M be a C-perfect matching on G. We construct a reporting strategy from M as follows. Consider any $x \in C$, $x' \in A$ such that $\{x, x'\} \in M \subseteq E$. When the true state is either x or x', the reporting strategy reveals the edge $\{x, x'\}$. That is, the sender reveals the k facts $x_K = x'_K$ that x and x' have in common, for some $K \subseteq \{1, ..., n\}$ and in addition reports that the state is either x or x'. She also chooses $d = a$. Since x and x' are equally likely, such a message and decision satisfies the deniability constraint (2). Since M is a C-perfect matching, this describes the sender’s report for each $x \in A \cup C$ that is matched by M.

For $x \in A$ that are unmatched by M, assume the sender reveals all aspects of x and chooses $d = a$. She also fully reveals all aspects of every $x \in R$ and chooses $d = r$. It is immediate that (2) is satisfied in both these cases. Since $d = a$ if and only if $x \in A \cup C$, we have a reporting strategy that implements the sender’s ideal decision rule while meeting the deniability constraint (2).

Step 2 (‘only if’). Suppose σ is reporting strategy that implements the sender’s ideal decision rule. Pick any $S \subseteq C$ and $x \in S$. To satisfy (2) any $m \in \textit{supp}\ \sigma(x)$ must also belong to $\textit{supp}\ \sigma(x')$ for some $x' \in N(\{x\})$. For each $x \in S$, define $\pi(x) \subseteq N(\{x\}) \subseteq A$ as follows: $x' \in \pi(x) \Leftrightarrow \textit{supp}\ \sigma(x) \cap \textit{supp}\ \sigma(x') \neq \emptyset$. Let $\pi(S) = \cup_{x \in S} \pi(x)$ and $\textit{supp}\ \sigma(S) = \cup_{x \in S} \textit{supp}\ \sigma(x)$. Note that $\pi(S) \subseteq N(S)$ and further that

$$\frac{1}{2} \leq \Pr[A \mid \textit{supp}\ \sigma(S), a] \leq \frac{|\pi(S)|}{|\pi(S)| + |S|}$$

The first inequality follows from using the law of iterated expectations and noting that for each $m \in \textit{supp}\ \sigma(S)$, (2) must hold. The second inequality follows from noting (i) that all states in $A \cup C$ are equiprobable, (ii) that any $x' \in \pi(S)$ may send $m \notin \textit{supp}\ \sigma(S)$ with strictly positive probability, and (iii) that some $x \notin S$ may send $m \in \textit{supp}\ \sigma(S)$ with strictly positive probability. This implies $|S| \leq |\pi(S)| \leq N(S)$, establishing (HC).■

Figure 1 provides an instance of Proposition 1 with $n = 2$, $k = 1$, $X_1 = \{0, 1, 2\}$, $X_2 = \{0, 1\}$ and uniform priors.[3] It is easy to see from the figure that (HC) holds. Panel (b) depicts the graph G as well as a C-perfect matching M (see marked edges). The strategy associated with this matching can be described as follows. The sender chooses $d = a$ by revealing $x_2 = 1$ when the state is either $(0, 1)$ or $(2, 1)$, and by revealing $x_1 = 1$ when the state is either $(1, 0)$ or $(1, 1)$. In state $(2, 0)$ she chooses $d = a$, while in $(0, 0)$ she chooses $d = r$, in both cases by revealing the values of both aspects. Since priors are uniform, (2) is met in all cases. Since $d = a$ for all $x \in A \cup C$ the sender obtains her ideal payoffs in each state.

In what follows we use the same graph theoretic approach to generalize Proposition 1 and cover non-uniform priors and general value functions u and v. We also characterize the sender's optimal reporting strategies when a C-perfect matching does not exist. But the core intuition of our approach is contained in the simple environment covered by Proposition 1. For the sender to obtain her ideal payoffs, each element of $x \in C$ must be pooled, or matched, with some element $x' \in A$. Condition (HC) states that there is enough diversity in the ways such pooling can occur for subversion to be possible. Proposition 1 shows that this condition is necessary and sufficient for the sender to be able to implement her ideal decision rule.

Notice that for the unconstrained persuasion problem with $k = 0$, every $x \in C$ can be pooled (or matched) with any $x' \in A$, i.e., G is a complete bipartite graph. In this case (HC) reduces to requiring $|C| \leq |A|$. At the other extreme where $k = n$, G is completely disconnected and $N(S)$ is empty for each x, so that (HC) fails and subversion is impossible. Condition (HC) keeps track of what is necessary and sufficient for subversion for each possible k. Since the edge set E corresponding to a given k is subset of the edge set corresponding to any $k' < k$, implementing her own ideal rule becomes more difficult for the sender when she is required to provide more supporting facts.

We now generalize Proposition 1. Fix the sender's problem $\mathcal{P}$ and recall the graph $G = \{A \cup C, E\}$. To each vertex $x \in A \cup C$ assign a weight $w(x) = \|p(x)u(x)\| > 0$. Unless otherwise specified, we will assume $w(x)$ is rational for all $x \in A \cup C$ throughout what follows. Using these vertex weights w, we construct an auxiliary 'cloned' graph ${}_wG(\mathcal{P}) = \{{}_wA \cup {}_wC, {}_wE\}$ as follows.

Let $h = L^{-1}$, where L denotes the least common multiple of the denominators of the (reduced fractions) $\{w(x)\}_{x \in A \cup C}$. For each $x \in A \cup C$ create $w(x)/h$ "clones" of x, each denoted by ${}^i_c x$, $i = 1, ..., w(x)/h$. Let ${}_wA$ be the set of all clones of all $x \in A$ and ${}_wC$ be the set of all clones of all $x \in C$. Construct ${}_wE$ according to the following rule: $\{{}^i_w x, {}^j_w x'\} \in {}_wE$, $i = 1, ..., w(x)/h$, $j \in 1, ..., w(x')/h$, if and only if $\{x, x'\} \in E$. Notice that when u coincides with the balance of

[3]The sets A, C and R are colored blue, yellow and red, respectively, in al our figures.

probabilities standard and all $x \in A \cup C$ are equiprobable, ${}_wG$ coincides with G. We are ready for the generalization of Proposition 1.

Proposition 2 *Assume $w(x)$ is rational for all $x \in A \cup C$. A subversive reporting strategy exists iff*

$$E[u \mid S \cup N(S)] \geq 0 \text{ for all } S \subseteq C. \tag{HC$_w$}$$

Proof.

Step 1 ('if'). We show below that when (HC$_w$) obtains on G, (HC) obtains on ${}_wG$. For now, suppose the latter condition holds so that by Hall's theorem a C-perfect matching exists on ${}_wG$. Let ${}_wM$ be this matching. Construct a subversive reporting strategy from ${}_wM$, as follows.

Consider any $x \in C$, $x' \in A$ such that $\{x, x'\} \in E$ and $\{{}^i_wx, {}^j_wx'\} \in {}_wM$ for some $i = 1, ..., w(x)/h$, $j \in 1, ..., w(x')/h$. Let $\mu(\{{}^i_wx, {}^j_wx'\})$ be a unique message that identifies the particular match $\{{}^i_wx, {}^j_wx'\} \in {}_wM$. The sender's report m consists of revealing the edge $\{x, x'\}$ as well as the message $\mu(\{{}^i_px, {}^j_px'\})$. She sends report m with probability $h/w(x)$ when the realized state is x, and with probability $h/w(x')$ when the realized state is x'. She also chooses $d = a$ after such a report. Note that

$$\begin{aligned}
\Pr[A \mid m, a] &= \frac{\frac{h}{w(x')}p(x')}{\frac{h}{w(x')}p(x') + \frac{h}{w(x)}p(x)} = -\frac{u(x)}{u(x') - u(x)} \\
&= 1 - \Pr[C \mid m, a]
\end{aligned}$$

implying

$$\mathbb{E}[u \mid m, a] = \Pr[A \mid m, a]\, u(x') + \Pr[C \mid m, a]\, u(x) = 0$$

so that the deniability constraint (1) is met.

Since ${}_wM$ is a C-perfect matching, for each $x \in C$ the the probabilities of such randomized reports add up to one. For $x' \in A$, these probabilities may add up to some number $q_{x'} < 1$ less than one, since some ${}^j_wx'$ may not be matched under ${}_wM$. In these cases, x' is perfectly revealed with the remaining probability $1 - q_{x'}$, and the sender chooses $d = a$. The sender also perfectly reveals all $x \in R$ and chooses $d = r$. In both cases, (1) is met. Since $d = a$ if and only if $x \in A \cup C$, this reporting strategy is subversive.

It remains to show (HC$_w$) on G implies (HC) on ${}_wG$. Pick a subset ${}_wS \subseteq {}_wC$. Construct $S \subseteq C$ as follows: $x \in S$ if ${}^i_wx \in {}_wS$ for some $i = 1, ..., w(x)/h$. Thus, S contains only those elements of

C for which at least one clone is contained in ${}_wS$. Let ${}_wS' = \cup_{x\in S} \cup_i \left\{{}^i_w x\right\} \subseteq {}_wC$ be the set of all clones of all elements of S. Note that for any $x \in C$, $N(\{{}^i_w x\}) = N(\{{}^j_w x\})$ for every pair of clones ${}^i_w x$ and ${}^j_w x$ of x. This implies

$$\begin{aligned} |N({}_wS)| &= \left|N({}_wS')\right| \\ &= \sum_{x'\in N(S)} \frac{w(x')}{h} \geq \sum_{x\in S} \frac{w(x)}{h} \\ &= \left|{}_wS'\right| \geq |{}_wS|, \end{aligned}$$

where the inequality in the second line follows from (HC$_w$) applied to ${}_wS'$ and the inequality in the third line follows from noting ${}_wS \subseteq {}_wS'$.

Step 2 ('only if'**).**

Suppose there exists a subversive reporting strategy σ. Pick $S \subseteq C$ and using the law of iterated expectations note that

$$\begin{aligned} E[u \mid S \cup N(S), a] = \quad & \Pr[S \cup \pi(S) \mid S \cup N(S), a]\mathbb{E}[u \mid S \cup \pi(S), a] \\ & + \Pr[N(S) \setminus \pi(S) \mid S \cup N(S), a]\mathbb{E}[u \mid N(S) \setminus \pi(S), a] \end{aligned}$$

where the set $\pi(S)$ has been defined in step 2 of the proof of Proposition 1. Since $u(x) > 0$ for all $x \in N(S) \setminus \pi(S) \subseteq A$, we must have

$$\mathbb{E}[u \mid S \cup N(S), a] \geq \Pr[S \cup \pi(S) \mid S \cup N(S), a]\mathbb{E}[u \mid S \cup \pi(S), a]$$

But, using the law of iterated expectations again,

$$\mathbb{E}[u \mid S \cup \pi(S), a] = \sum_{m \in supp\ \sigma(S\cup\pi(S))} \Pr[m \mid S \cup \pi(S), a]\mathbb{E}[u \mid m, a] \geq 0,$$

where $supp\ \sigma(S \cup \pi(S)) = \cup_{x\in S\cup\pi(S)} supp\ \sigma(x)$; and the inequality follows from the fact that σ is a subversive reporting strategy that meets (1) for all $m \in supp\ \sigma(S \cup \pi(S))$. This establishes (HC$_w$) on G.■

The cloned graph ${}_wG$, when viewed as the primitive, describes an environment similar to those covered by Proposition 1. In particular, each vertex of ${}_wG$ can be thought to be equiprobable, with probability equal to h and the receiver uses the balance of probability standard (2) when evaluating the pooled message corresponding to an edge $\{{}^i_w x, {}^j_w x'\} \in {}_wM \subseteq {}_wE$, that is sent with positive probability both states x and x'. Proposition 2 links Hall's condition (HC) when applied to ${}_wG$ to an equivalent condition (HC$_w$) on the applied to the actual problem. It makes precise the sense in which assuming uniform priors and a balance of probabilities standard are without

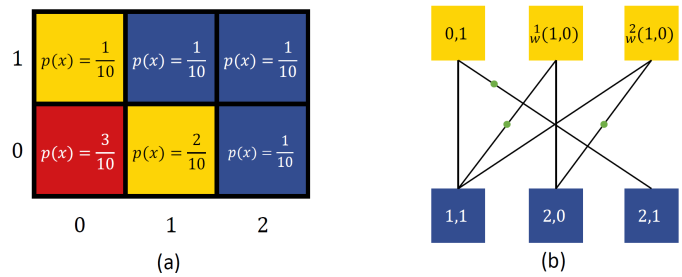

Figure 2: A perfect matching on ${}_wG$.

loss of generality, even if actual receiver preferences are more general and even when priors display correlation across aspects.

Figure 2 provides an example of Proposition 2 with $n = 2$, $k = 1$, $X_1 = \{0, 1, 2\}$, $X_2 = \{0, 1\}$. Panel (a) depicts the sets A, C and R as well as the priors p that display correlation. We suppose $u = +1$ for all $x \in A$ with $u = -1$ otherwise, so that the receiver uses the balance of probabilities standard (2). Panel (b) depicts the cloned graph ${}_wG$ as well as a perfect matching ${}_wM$ (see marked edges).

Using ${}_wM$ we construct the sender's subversive reporting strategy. When $x = (1, 1)$ the sender reveals $x_1 = 1$, and when $x = (2, 0)$ she reveals $x_2 = 0$. In both cases she chooses $d = a$. She also chooses $d = a$ when $x = (1, 0)$, either revealing $x_1 = 1$ or revealing $x_2 = 0$, each with probability 1/2. Because of this randomization, (2) is met when the receiver sees $x_1 = 1$ and $d = a$ and also when he sees $x_2 = 0$ and $d = a$. The sender also chooses $d = a$ and reveals $x_2 = 1$ when $x = (0, 1)$ and when $x = (2, 1)$. She chooses $d = r$ when $x = (0, 0)$. The deniability constraint is met in all these cases as well. Since the sender obtains her ideal decision in every state, this reporting strategy is subversive.

To obtain Proposition 2 we assume each $w(x)$ is rational. This restriction is not important for the 'only if' part, as an inspection of the proof reveals. Since rationals are dense in the reals, it is also not a substantive restriction for the 'if' part as the following corollary demonstrates.

Corollary 1 *A subversive reporting strategy exists if* $E[u \mid S \cup N(S)] > 0$ *for all* $S \subseteq C$.

Proof. For each $x \in X$, approximate $u(x)$ by $\widehat{u}(x) \in \mathbb{Q}$, so that $\widehat{u}(x) \leq u(x)$ and $\mathbb{E}[\widehat{u} \mid S \cup N(S)] > 0$ for all $S \subseteq C$, where the expectation is taken with respect to the prior p. Such an approximation exists by the density of rationals in the reals. For each $x \in X$ approximate $p(x)$ by $\widehat{p}(x) \in \mathbb{Q}$ such that $\widehat{p}(x) > p(x)$, and for each $x \notin C$ approximate $p(x)$ by $\widehat{p}(x) \in \mathbb{Q}$ such that

$\widehat{p}(x) < p(x)$, so that $\sum_{x\in X} \widehat{p}(x) = 1$ and $\mathbb{E}[\widehat{u} \mid S \cup N(S)] \geq 0$ where the last expectation is taken with respect to priors $\widehat{p}$. Applying Proposition 2 to the problem where u and p are replaced by $\widehat{u}$ and $\widehat{p}$ yields a C-perfect matching, and an associated subversive reporting strategy. Note that this matching in the new problem is still valid in the original problem since the approximations where chosen to ensure (1) is satisfied in the original problem for each valid match in the new problem.■

For a subversive reporting strategy, it is only important to specify the sender's ordinal ranking of the decisions in each state, i.e., to specify the sets $A \cup C$ and R. The following corollary of Proposition 2 describes simple cases where we can vary these sets and still ensure subversion is possible.

Corollary 2 *If a subversive reporting strategy exists for a problem* $\mathcal{P} = \langle \{A, C, R\}, \{p, u, v\}, n, k\rangle$ *then it also exists for any other problem* $\mathcal{P}' = \langle \{A', C', R'\}, \{p, u, v\}, n, k\rangle$ *satisfying* $A \subseteq A'$ *and* $C' \subseteq C$.

If (HC_w) is satisfied in $\mathcal{P}$ it is also satisfied in $\mathcal{P}'$. So subversion is easier in $\mathcal{P}'$ compared to $\mathcal{P}$. Given a subversive reporting strategy for $\mathcal{P}$, we can create one for $\mathcal{P}'$ simply by changing the sender's decision from $d = a$ to $d = r$ for all $x \in C \cap R'$, and conversely for all $x \in R \cap B'$, keeping everything else unchanged. Notice finally that the sender's cardinal payoff function v has not had a role to play for the results presented so far. It will have an important role in the next section where we characterize the sender's optimal reporting strategy when subversion is impossible.

2.3 Optimal reporting strategies

When (HC_w) fails, a C-perfect matching does not exist. We turn now to characterizing the sender's optimal reporting strategies in such situations. Since subversion is impossible, the sender must choose $d = r$ with strictly positive probability for some $x \in A \cup C$, ideally those which are least costly for her to give up on. The specification of the sender's cardinal payoffs v now becomes important. As shown below, we can adjust for the failure of (HC_w) by adding fictitious vertices and edges to ${}_wG$, taking the sender's cardinal utility into account by assigning suitable edge weights to each edge of the resulting graph. First we present a lemma that establishes a tight relationship between optimal reporting strategies and matching on a bipartite graph that has rational vertex weights.

Lemma 1 *Assume* $p(x)u(x)$ *is rational for all* $x \in A \cup C$. *For any reporting strategy* σ, *there exists a reporting strategy* σ' *such that (1) all of the messages are sent with rational probabilities, i.e.,* $\sigma'(x) \in \mathbb{Q}^{\mathcal{M}(x)}$ *for all* $x \in X$, *and (2) the sender's expected utility is weakly higher under* σ' *than under* σ.

Proof. See the Appendix.■

Using the lemma, our next result characterizes the sender's optimal strategy in a special case where the sender and the receiver both want to avoid mistakes and value all mistakes equally.

Proposition 3 *Assume (i) $v(x) = +1$ for $x \in A \cup C$ with $v(x) = -1$ otherwise, (ii) deniability corresponds to the balance of probability standard (2) and (iii) $p(x) \in \mathbb{Q}$ for all $x \in X$. The sender-optimal reporting strategy guarantees expected payoffs of* $\Pr(A) + \Pr(C) - \delta$ *to the sender and* $\Pr(A) - \Pr(C) + \delta$ *to the receiver, where*

$$\delta \equiv \max\left\{0, \max_{S \subseteq C} \left(-\mathbb{E}[u \mid S \cup N(S)]\right)\right\}$$

Proof. Note first that $\delta > 0$ iff (HC$_w$) fails. Because of Theorem 2, we can take $\delta > 0$ in order to identify an optimal reporting strategy σ. Furthermore, by lemma 1, it is without loss of generality to consider σ which uses rational probabilities only, i.e., for all x, $\sigma(x) \in \mathbb{Q}^{\mathcal{M}(x)}$.

Given σ, construct an auxiliary graph ${}_wG(\sigma) = \{{}_wA \cup {}_wC, {}_wE\}$ from G, as follows. To each $x \in A \cup C$ assign a weight $w(x) = p(x) > 0$. Let $h = L^{-1}$, where L denotes the least common multiple of the denominators of $\{w(x)\sigma(x)[m]\}_{x \in A \cup C, m \in \mathcal{M}}$. For each $x \in A \cup C$ create $w(x)/h$ clones of x, denoted by ${}_c^i x$, $i = 1, ..., w(x)/h$. Let ${}_wA$ be the set of all clones of all $x \in A$ and ${}_wC$ be the set of all clones of all $x \in C$. Construct ${}_wE$ according to the following rule: $\{{}_w^i x, {}_w^j x'\} \in {}_wE$, $i = 1, ..., w(x)/h$, $j \in 1, ..., w(x')/h$, if and only if $\{x, x'\} \in E$.

Note that σ induces a matching on this cloned graph. For any message $m \in supp_a\ \sigma$ we can match vertices ${}_w^i x \in {}_wC$, such that $m \in supp_a\ \sigma(x)$ with vertices $x' \in {}_wA$, such that $m \in supp_a\ \sigma(x')$. That is, we can match $\min\{p(x)\sigma(x)[m]\,h, p(x')\sigma(x')[m]\,h\}$ from each such pair of x and x'. Since $\sum_{x \in X} p(x)u(x)\sigma(x)[m] \geq 0$, there must be enough $x' \in {}_wA$ to match all of the $x \in {}_wC$. Note also that each match increases the sender's utility by h.

The Konig-Ore formula (Lovasz and Plummer, 1986, Theorem 1.3.1) states that the maximal matching on ${}_wG$ matches all but $\max_{{}_wS \subseteq {}_wC} [|{}_wS| - |N({}_wS)|]$ points in ${}_wC$. Thus, if σ is optimal it must leave exactly that many vertices in ${}_wC$ unmatched. Since for any $x \in C$, every pair of clones ${}_w^i x$ and ${}_w^j x$ of x has the same neighbor set, i.e., $N(\{{}_w^i x\}) = N(\{{}_w^j x\})$, we have that all clones of any $x \in C$ must be included in the set maximizing $\max_{{}_wS \subseteq {}_wC} [|{}_wS| - |N({}_wS)|]$. Thus, the number of unmatched points in ${}_wC$ under the optimal strategy is

$$\begin{aligned}\max_{{}_wS\subseteq{}_wC}\left[|{}_wS|-|N\left({}_wS\right)|\right] &= \max_{S\subseteq C}\left[\sum_{x\in S}\left|\frac{w(x)}{h}\right|-\sum_{x'\in N(S)}\left|\frac{w(x')}{h}\right|\right]\\ &= \frac{1}{h}\max_{S\subseteq C}\left[\sum_{x\in S}w(x)-\sum_{x'\in N(S)}w(x')\right]\\ &= \frac{1}{h}\max_{S\subseteq C}\left[-\mathbb{E}[u\mid S\cup N\left(S\right)]\right].\end{aligned}$$

Since each unmatched element of ${}_wC$ has utility loss to the sender of h, while she gets her preferred action for all $x\in A$ and cannot do better than choosing $d=a$ for all $x\in A\cup C$, the sender's expected utility from the optimal strategy is:

$$\begin{aligned}&\Pr\left[A\right]+\Pr\left[C\right]+(-h)\frac{1}{h}\max_{S\subseteq C}\left[-\mathbb{E}[u\mid S\cup N\left(S\right)]\right]\\ =\;&\Pr\left[A\right]+\Pr\left[C\right]-\delta.\end{aligned}$$

Since the receiver prefers $d=r$ for $x\in C$, the receiver's expected utility from the optimal strategy is also easily seen to be $\Pr\left[A\right]-\Pr\left[C\right]+\delta$. ■

To see why Proposition 3 obtains consider again the environment of Proposition 1 where all $x\in A\cup C$ are equiprobable and the deniability constraint corresponds to the balance of probabilities standard (2). If (HC$_w$) (equivalently, (HC)) fails in G, we can add δ additional "fictitious" vertices to A and connect every such fictitious vertex to every vertex in C. The resulting graph must satisfy (HC), implying there exists a C-perfect matching M_δ on it. A *maximal matching* can be derived from M_δ by keeping unmatched all vertices $x\in C$ that have a fictitious match according to M_δ, keeping all other matches unchanged. The proof of Proposition 3 completes this argument by using lemma 1 and creating the cloned graph ${}_wG(\sigma)$ on which the sender's optimal reporting strategy corresponds to a matching. It also allows for arbitrary rational priors.

The balance of probabilities standard (2) is imposed for Proposition 3 in order to make sure the sender's and receiver's payoffs are each proportional to the number of matches. A similar approach will work if we allow more general u and v, as long as we add a suitable edge weighting function to an appropriately constructed bipartite graph. We show this next.

Fix sender's problem of selecting facts, $\mathcal{P}$. Let ${}_\eta G=\{{}_wA\cup{}_wC,{}_wE,\eta\}$ denote an edge-weighted graph, where $\eta:{}_wE\rightarrow\mathbb{R}$ is the edge weight function. The graph ${}_\eta G$ is constructed similarly to earlier graphs. To each $x\in A\cup C$ assign a weight $w(x)=\|p\left(x\right)u\left(x\right)\|\in\mathbb{Q}$ and let $h=L^{-1}$ where L is the least common multiple of the denominators of $\{w(x)\}_{x\in A\cup C}$. For each $x\in A\cup C$ create $w(x)/h$ clones of x, denoted by ${}_w^ix$, $i=1,...,w(x)/h$. Let ${}_wA$ be the set of all clones of all $x\in A$ and ${}_wC$

be the set of all clones of all $x \in C$. Construct ${}_wE$ according to the following rule: $\{{}^i_wx, {}^j_wx'\} \in {}_wE$, $i = 1, ..., w(x)/h$, $j \in 1, ..., w(x')/h$, if and only if $x \in C$ and $x' \in A$ and $x_K = x'_K$ for some $K \subseteq \{1, ..., n\}$ with $|K| = k$. The edge-weight function is given by $\eta\left(\{{}^i_wx, {}^j_wx'\}\right) = hp(x)\, v(x)$ for $x \in C$.

Let the weight of a matching on ${}_\eta G$ be the sum of the weights of the edges contained in the matching. We show that the maximum weight matching on graph ${}_\eta G$ is an optimal reporting strategy for the sender. The total edge-weight of the maximum matching on a graph G is called the matching number and we denote it $\nu^*(G)$. Let a maximum weight matching on ${}_\eta G$ be denoted by ${}_\eta M \subset {}_wG$.

Proposition 4 *The sender-optimal strategy for problem $\mathcal{P}$ gives the sender a payoff of $\nu^*(G(\mathcal{P})) + \sum_{x \in A} p(x)v(x)$.*

Proof.

The sender can guarantee the first-best outcome on all $x \in A \cup R$, so the problem reduces to determining what the sender can get on the set C. Let σ be a sender-optimal strategy for $\mathcal{P}$. Note that by lemma **??**, it is without loss of generality to consider σ for which $\sigma(x) \in \mathbb{Q}^{\mathcal{M}(x)}$ for all x.

Given σ,define ${}_\sigma G$ as a edge-weighted graph that is a 'finer' version of ${}_\eta G$. In particular, let $h' = 1/L'$, where L' denotes the least common multiple of the denominators of $\{\sigma(x)[m]\}_{x \in A \cup C, m \in \mathcal{M}}$. For each ${}^i_wx \in {}_wA \cup {}_wC$ create h' clones of ${}^i_wx$, denoted by ${}^{i,j}_\sigma x$, $i = 1, ..., w(x)/h$, $j = 1, ..., 1/h'$. Let ${}_\sigma A$ be the set of all clones of all ${}^i_wx \in {}_wA$ and ${}_\sigma C$ be the set of all clones of all ${}^i_wx \in {}_wC$. Construct ${}_\sigma E$ as follows: $\{{}^{i,j}_\sigma x, {}^{k,l}_\sigma x'\} \in {}_\sigma E$, for all i,j,k,l if and only if $x_K = x'_K$ for some $K \subseteq \{1, ..., n\}$ with $|K| = k$. Adjust the edge weights to equal $hp(x)\, v(x)\, h'$, $x \in C$.

Note that σ induces a matching on graph ${}_\sigma G$. For any message $m \in supp_a\ \sigma$ we can match vertices ${}^{i,j}_\sigma x \in {}_\sigma C$, $m \in supp_a\ \sigma(x)$, with vertices ${}^{k,l}_\sigma x' \in {}_\sigma A$, $m \in supp_a\ \sigma(x')$. There must be enough ${}^{k,l}_\sigma x' \in {}_\sigma A$ to match all of the ${}^{i,j}_\sigma x \in {}_\sigma C$ since $\sum_{x \in X} p(x)\, u(x)\, \sigma(x)[m] \geq 0$. Denote this matching by ${}_\sigma M$. A strategy that maximizes the sender's payoff must be a maximum weight matching on the graph ${}_\sigma G$, since any matching on this graph is a feasible strategy satisfying (1).

It remains to show that $\nu^*({}_\sigma G) = \nu^*({}_\eta G)$. We show first that $\nu^*({}_\sigma G) \geq \nu^*({}_\eta G)$. Note that any matching on ${}_\eta G$ induces a matching on the finer graph ${}_\sigma G$ since we can just take every edge in the matching $\{{}^i_wx, {}^j_wx'\} \in {}_\eta M$ and match the extra h' copies of ${}^i_wx$ to the extra h' copies of ${}^j_wx'$.

It follows that

$$\begin{aligned}
\nu^* \left({}_{\eta}G \right) &= \sum_{{}^{i}_{w}x \in {}_{w}C \cap {}_{\eta}M} hp\left(x\right) v\left(x\right) \\
&= \sum_{{}^{i}_{w}x \in {}_{w}C \cap {}_{\eta}M} hp\left(x\right) v\left(x\right) \frac{h'}{h'} \\
&= \sum_{{}^{i}_{w}x \in {}_{w}C \cap {}_{\eta}M} \sum_{j=1}^{1/h'} hp\left(x\right) v\left(x\right) h' \\
&\leq \nu^* \left({}_{\sigma}G \right),
\end{aligned}$$

where the inequality follows since the third line is an expression for the sum of edge-weights of a matching on ${}_{\sigma}G$.

To show the reverse inequality, $\nu^* \left({}_{\sigma}G \right) \leq \nu^* \left({}_{\eta}G \right)$, we want to construct a matching on ${}_{\eta}G$ from the maximal matching on ${}_{\sigma}G$. Note that we can assume without loss of generality that ${}_{\sigma}M$ does not contain cycles of clones, e.g., that there do not exist $\{{}^{i,j}_{\sigma}x_C, {}^{k,l}_{\sigma}x_A\}$, $\{{}^{k,l'}_{\sigma}x_A, {}^{i',j'}_{\sigma}x'_C,\}$, $\{{}^{i',j''}_{\sigma}x'_C, {}^{k',l''}_{\sigma}x'_A\}$ and $\{{}^{k',l'''}_{\sigma}x_A, {}^{i,j'''}_{\sigma}x_C\} \in {}_{\sigma}M$, since then we could generate an equal-weight matching by replacing the links $\{{}^{k,l'}_{\sigma}x_A, {}^{i',j'}_{\sigma}x'_C,\}$ and $\{{}^{k',l'''}_{\sigma}x_A, {}^{i,j'''}_{\sigma}x_C\}$ with $\{{}^{k,l'}_{\sigma}x_A, {}^{i,j'''}_{\sigma}x_C\}$ and $\{{}^{i',j'}_{\sigma}x'_C, {}^{k',l'''}_{\sigma}x_A\}$. The weight of the matching is unchanged since the weights of each edge is only a function of $x \in C$.

Consider each ${}^{i}_{w}x \in {}_{w}C$ and its $1/h'$ clones ${}^{i,j}_{\sigma}x \in {}_{\sigma}C$. We want to show that there exists a matching on ${}_{\sigma}G$ with the same weight as ${}_{\sigma}M$, called ${}_{\sigma}M'$, such that $\{{}^{i,j}_{\sigma}x_C, {}^{k,l}_{\sigma}x_A\} \in {}_{\sigma}M'$ implies that $\{{}^{i,j'}_{\sigma}x_C, {}^{k,l'}_{\sigma}x_A\} \in {}_{\sigma}M'$ for all j',$k' \in 1, ..., 1/h'$, i.e., that all these σ-clones are matched together. This allows us to construct a matching on ${}_{\eta}G$, since where all ${}^{i}_{w}x_C$ and ${}^{k}_{w}x_A$ are matched which each other. Since cycles have been ruled out, the only remaining possibility is that we have a "clone path", i.e., that $\{{}^{i,j}_{\sigma}x_C, {}^{k,l}_{\sigma}x_A\}, \{{}^{k,l'}_{\sigma}x_A, {}^{i',j'}_{\sigma}x'_C,\}, \{{}^{i',j''}_{\sigma}x'_C, {}^{k',l''}_{\sigma}x'_A\}, \{{}^{k',l'''}_{\sigma}x_A, {}^{i'',j'''}_{\sigma}x''_C\} \in {}_{\sigma}M$ and that ${}^{i,j^*}_{\sigma}x_C, {}^{i'',j\hat{}}_{\sigma}x''_C$ are unmatched for some j^* and $j\hat{}$ (the path can be longer than this, but either it's a cycle or it leads to some unmatched end points). This can only be optimal if $v\left(x_C\right) = v\left(x''_C\right)$, since if $v\left(x_C\right) > v\left(x''_C\right)$ a matching which included ${}^{i,j^*}_{\sigma}x_C$ at the expense of ${}^{i'',j'''}_{\sigma}x''_C$ would have given a higher weight. Given the indifference, in constructing our matching ${}_{\sigma}M'$ we can replace $\{{}^{k,l'}_{\sigma}x_A, {}^{i',j'}_{\sigma}x'_C,\}$ and $\{{}^{k',l'''}_{\sigma}x_A, {}^{i'',j'''}_{\sigma}x''_C\}$ with $\{{}^{i,j^*}_{\sigma}x_C, {}^{k,l'}_{\sigma}x_A\}$ and $\{{}^{i',j''}_{\sigma}x'_C, {}^{k',l'''}_{\sigma}x_A\}$. Replace all such clone paths and note that the resulting matching ${}_{\sigma}M'$ has the desired property that all σ-clones are matched together. Note that the induced matching on ${}_{\eta}G$ is one possible matching and thus $\nu^* \left({}_{\sigma}G \right) \leq \nu^* \left({}_{\eta}G \right)$. This establishes the result.■

This concludes our discussion of the problem of selecting persuasive facts under commitment. The graph theoretic techniques that we employ allow us to establish results of some generality that go beyond the geometric structure of our particular formulation. In our problem, the sender can match (or pool) $x \in C$ with $x' \in A$ if x and x' are 'similar' in that they share the values of at least k aspects. These k shared facts define an edge $\{x, x'\}$. But if we take the edges as the primitives, then the resulting graphs allow many alternative notions of what states the expert can pool, i.e., what counts as a persuasive argument.

For instance, we could assume that if $x \in C$ and $x' \in A$ are within a given distance then the sender does not need to reveal any facts and k can equal zero. This captures the idea that nearby points require no supporting facts or special proof and their similarity is obvious to the receiver. Pooling more distant points may require more supporting facts. Once we construct a graph to capture such a notion of similarity our results will apply unchanged. So our approach promises a characterization of persuasive arguments across a rich variety of environments. We postpone such generalizations for future research. In the next section we turn to an analysis of what happens without commitment, situations where the geometric structure of our particular notion of similarity has a greater role to play.

3 The case without commitment

To Be Added

4 Conclusion

To Be Added

5 Appendix

Proof of Lemma 1. For a reporting strategy σ and a message m, let $\sum_{x \in X} p(x) u(x) \sigma(x)[m] := b_m$. for an optimal reporting strategy we must have: (1) for any $x \in X$, $\sum_{m \in \mathcal{M}} \sigma(x)[m] = 1$, and (2) for any $m \in supp_a\ \sigma$, we have $b_m \geq 0$ and (3) for any $m \in supp_r\ \sigma$, $b_m \leq 0$.

Note that a reporting strategy can always perfectly reveal points in the set R and A where the sender and receiver agree and so any strategy that makes a mistake on these subsets of X can be improved upon. Accordingly, restrict attention to strategies where for $x \in A$, $supp_a\ \sigma(x) = supp\ \sigma(x)$ and for $x \in R$, $supp_r\ \sigma(x) = supp\ \sigma(x)$. In particular, replace the messaging strategy by one that perfectly reveals each state in R with probability 1 (hence, it is rational) and perfectly reveals all $x \in A$ if which are not pooled with any $x \in C$, or which when pooled with $x \in C$ lead to the reject decision. This means that any $m \in supp_r\ \sigma$ should not be sent with positive probability by any $x \in A$ since simply revealing that x would be improve the sender's payoffs. Thus, for any $m \in supp_r\ \sigma$, $b_m < 0$, i.e., the receiver would earn negative payoffs from $d = a$ in those states.

Furthermore, we can remove redundant messages (e.g., different messages that reveal the same state or the same set of states) by pooling them together into a single message. This ensures that we have a finite number of messages in the support of σ. Finally, we can replace any $m \in \cup_{x \in C} supp_r\ \sigma(x)$ with a single message m_r which indicates that the state is in C and should be rejected.

Let $\sigma^0 = \sigma$ and $M^+_{\sigma^0}$ be the set of messages $m \in supp_a\ \sigma^0$ such that $b_m > 0$. For each $x_1 \in A$ and $m_1 \in supp_a\ \sigma^0(x_1) \cap M^+_{\sigma^0}$, make a rational approximation of $\sigma^0(x_1)$, called $\sigma^1(x_1)$, such that $\sigma^1(x_1)[m_1] < \sigma^0(x_1)[m_1]$ and $\sigma^1(x_1)[m] > \sigma^0(x_1)[m]$ for all $m \neq m_1$, $m \in supp\ \sigma^0(x_1)$. For all $x_0 \in C$ and $m_0 \in supp_a\ \sigma^0(x_0) \cap M^+_{\sigma^0}$, make a rational approximation of $\sigma^0(x_0)$, called $\sigma^1(x_0)$, such that $\sigma^1(x_0)[m_0] > \sigma^0(x_0)[m_0]$ and $\sigma^1(x_0)[m] < \sigma^0(x_0)[m]$ for all $m \neq m_1$, $m \in supp\ \sigma^0(x_0)$. For all other x and m, let $\sigma^1(x)[m] = \sigma^0(x)[m]$. Ensure that these approximations maintain that $b^1_m := \sum_{x \in X} p(x) u(x) \sigma^1(x)[m] > 0$ for all $m \in M^+_{\sigma^0}$. This is possible by the density of rationals and since $b_{m_1} > 0$. These rational approximations can only make the sender better off (strictly so when $x_0 \in C$, and $\sigma^1(x_0)[m] < \sigma^0(x_0)[m]$ for m that leads to rejection). These approximations may result in messages $m \notin M^+_{\sigma^0}$ having the property that $b^1_m > 0$. Let $M^+_{\sigma^1}$ be the set of messages in the support of σ^1 such that $b^1_m > 0$, and observe that $M^+_{\sigma^1} \supseteq M^+_{\sigma^0}$. We iterate on this process, until no new messages are added in this way. Since the number of messages in the support σ is finite, this process converges.

Call the resulting strategy σ' and let $b'_m = \sum_{x \in X} p(x) u(x) \sigma'(x)[m]$. Note that for any m such that $b'_m > 0$, and any x such that $m \in supp\ \sigma'(x)$ we have that $\sigma'(x)$ is rational. Consider an m_1 such that $b'_{m_1} = 0$ and that $\sigma'(x_1)[m_1] \notin \mathbb{Q}$ for some x_1. We must have that there exists some set $X_1 \subseteq X$ such that $m_1 \in supp\ \sigma'(x)$ for $x \in X_1$ and that for all other messages m in the support of some $x \in X_1$ we have $b_m \leq 0$.

Suppose first that $b_{m_1} = 0$ for all $m_1 \in \cup_{x \in X_1} supp\ \sigma'(x)$. In this case, we can write the reporting strategy $\sigma'(x)[m]$ for $x \in X_1$ using rationally independent set of real numbers, i.e., each $\sigma'(x)[m]$ is a linear combination with rational coefficients of these numbers and that the real numbers in this set cannot be expressed as a rational combinations of each other. Now approximate each irrational in this rationally independent set by a rational number. Note that since these irrationals were rationally independent, after we approximate and replace $\sigma'(x)[m]$ by the same linear combination as before except using the rational approximation, we still have $\sum_{m \in \mathcal{M}} \sigma'(x)[m] = 1$ for all $x \in X_1$ and $\sum_{x \in X_1} w(x) \sigma'(x)[m_1] = 0$. We now have that $b_m \in \mathbb{Q}$ for all $m \in supp_a\ \sigma$.

Suppose next that $m_r \in \cup_{x \in X_1} supp\ \sigma'(x)$, with $b_r < 0$ (by the definition of b_r). We note that

$b_r \in \mathbb{Q}$, since

$$\begin{aligned} b_r &= \sum_{x \in C} p(x) u(x) \left(1 - \sum_{m \in supp_a \ \sigma} \sigma(x)[m]\right) \\ &= \sum_{x \in C} p(x) u(x) - \sum_{x \in C} \sum_{m \in supp_a \ \sigma} p(x) u(x) \sigma(x)[m], \end{aligned}$$

is the difference of two rational numbers. To show that the second term is rational, observe that

$$\begin{aligned} \sum_{m \in supp_a \ \sigma} b_m &= \sum_{m \in supp_a \ \sigma} \left(\sum_{x \in A} p(x) u(x) \sigma(x)[m] + \sum_{x \in C} p(x) u(x) \sigma(x)[m]\right) \\ &= \sum_{m \in supp_a \ \sigma} \sum_{x \in A} p(x) u(x) \sigma(x)[m] + \sum_{m \in supp_a \ \sigma} \sum_{x \in C} p(x) u(x) \sigma(x)[m] \\ &= \sum_{x \in A} p(x) u(x) + \sum_{x \in C} \sum_{m \in supp_a \ \sigma} p(x) u(x) \sigma(x)[m], \end{aligned}$$

so that

$$\sum_{x \in C} \sum_{m \in supp_a \ \sigma} p(x) u(x) \sigma(x)[m] = \sum_{m \in supp_a \ \sigma} b_m - \sum_{x \in A} p(x) u(x) \in \mathbb{Q}.$$

Once again we can write $\sigma(x)[m]$ for $x \in X_1$ using rational linear combinations of a rationally-independent set of reals (the set may need to include b_r). Since we can have an irrational number come in with both positive and negative rational coefficients in the representation of $\sigma(x)[m_r]$, we need to decide whether to approximate that irrational from above or below. This is determined by the sender's preferences. In particular we can have $x_1, x_2, ..., x_j \in C$ such that some irrational number, say n_m, is in the linear combination of each $\sigma(x_i)[m_r]$ for $i = 1, ..., j$. That is, $\sigma(x_i)[m_r] = r_1^i n_1 + ... + r_k^i n_k$, for rational coefficients $r_1, ..., r_k$ and real numbers $n_1, ..., n_k$. For each such n_m we compute $\sum_{i=1}^{j} v(x_i) r_m^i$, the derivative with respect to n_m of the loss to the sender from taking the reject action in states $X_1 \cap C$. We approximate the number n_m from below if this sum is positive, and from above if the sum is negative. We do so for each irrational in the rationally independent set and note that this improves the sender's payoff. We have now constructed a rational reporting strategy for the sender which does at least as well as any original strategy. ■

References

[1] Antic, Nemanja, Archishman Chakraborty and Rick Harbaugh. 2022. "Subversive Conversations," *Working Paper*.

[2] Battaglini, Marco. 2002. "Multiple Referrals and Multidimensional Cheap Talk," *Econometrica*, 70(4): 1379–1401.

[3] Chakraborty, Archishman and Rick Harbaugh. 2010. "Persuasion by Cheap Talk," *American Economic Review*, 100(5): 2361–2382.

[4] Crawford, Vincent P. and Joel Sobel, 1982. "Strategic Information Transmission," *Econometrica*, 1431–1451.

[5] Duan, R., and Su, H. H. 2012. "A scaling algorithm for maximum weight matching in bipartite graphs." In Proceedings of the twenty-third annual ACM-SIAM symposium on Discrete Algorithms: 1413-1424.

[6] Dziuda, Wioletta. 2011. "Strategic Argumentation," *Journal of Economic Theory*, 146(4): 1362-1397.

[7] Farrell, Joseph and Robert Gibbons. 1989. "Cheap Talk with Two Audiences," *American Economic Review*, 79(5): 1214–1223.

[8] Fishman, Michael J. and Kathleen M. Hagerty. 1989. "Disclosure decisions by firms and the competition for price efficiency," *The Journal of Finance* 44(3): 633-646.

[9] Glazer, Jacob and Ariel Rubinstein. 2004. "On Optimal Rules of Persuasion," *Econometrica*, 72(6): 1715–36.

[10] Green, Jerry R., and Nancy L. Stokey. 2007. "A Two-Person Game of Information Transmission," *Journal of Economic Theory*, 135(1): 90–104.

[11] Hall, P., 1935. "On Representatives of Subsets," *Journal of the London Mathematical Society*, 10: 26–30.

[12] Kamenica, Emir and Matthew Gentzkow. 2011. "Bayesian Persuasion," *American Economic Review*, 101(6): 2590.

[13] Kolotilin, Anton, Corraro, Roberto and Alexander Wolitzky. 2022. "Persuasion with Non-Linear Preferences," *Working Paper.*

[14] Lipnowski, Elliot and Doron Ravid. 2020. "Cheap Talk with Transparent Motives," *Econometrica*, 88(4): 1631–1660.

[15] Lovász, László, and Michael D. Plummer. 2009. *Matching theory.* American Mathematical Society: 367.