Prophet Inequalities for Bayesian Persuasion

Niklas Hahn[1] , **Martin Hoefer**[1] , **Rann Smorodinsky**[2]
[1] Goethe University Frankfurt, Germany
[2] Technion, Israel
{nhahn, mhoefer}@em.uni-frankfurt.de, rann@ie.technion.ac.il

Abstract

We study an information-structure design problem (i.e., a Bayesian persuasion problem) in an online scenario. Inspired by the classic gambler's problem, consider a set of candidates who arrive sequentially and are evaluated by one agent (the *sender*). This agent learns the value from hiring the candidate to herself as well as the value to another agent, the *receiver*. The sender provides a signal to the receiver who, in turn, makes an irrevocable decision on whether or not to hire the candidate. A-priori, for each agent the distribution of valuation is independent across candidates but may not be identical. We design good online signaling schemes for the sender. To assess the performance, we compare the expected utility to that of an optimal offline scheme by a prophet sender who knows all candidate realizations in advance.

We show an optimal prophet inequality for online Bayesian persuasion, with a 1/2-approximation when the instance satisfies a "satisfactory-status-quo" assumption. Without this assumption there are instances without any finite approximation factor. We extend the results to combinatorial domains and obtain prophet inequalities for matching with multiple hires and multiple receivers.

1 Introduction

In many settings an informed agent wants to use private information in order to persuade other agents to take some action that he would benefit from. Consider a salesperson informed about the quality of the product who would like to maximize sales. Is full disclosure of the product quality to customers an optimal strategy? Perhaps revealing no information or revealing it partly would result in a higher sales volume.

The study of optimal information disclosure, known as *Bayesian persuasion*, has gained enormous attention in the recent decade. The canonical model is one where the informed agent (the *sender*) commits to some information disclosure (or *signaling*) scheme before learning the true state of nature. Once the state is realized, the appropriate signal is sent to other agents (the *receivers*) who, in turn, take an action which results in payoffs for the sender and the receivers. Applications abound and can be found in diverse areas such as online advertisement [Badanidiyuru *et al.*, 2018; Emek *et al.*, 2012; Arieli and Babichenko, 2019], security problems [Rabinovich *et al.*, 2015; Xu *et al.*, 2015, 2016], medical research [Kolotilin, 2015], financial-sector stress testing [Goldstein and Leitner, 2018], and voter coalition formation [Alonso and Câmara, 2016].

There are scenarios where this canonical model does not work well. For example, consider a routing problem where autonomous agents try to minimize their travel time over a network. Information about network congestion is available to a central planner who may share it with the agents in order to maximize the network's throughput. The planner receives the information gradually (road closures, broken traffic lights, traffic surges, etc.) and shares it gradually with the agents, who sequentially need to decide on the next edge to take in the network. In this case, we need to study *online* variants of the persuasion problem.

In this paper, we study the online version of the persuasion problem in a simple setting. Inspired by the classical gambler's problem, we consider a setting where n tasks arrive sequentially. A sender learns the completion value of each incoming task, for herself as well as for the receiver. Before the tasks arrive the sender commits to some signaling scheme and at each stage, depending on this policy and the actual information received, the receiver decides whether or not to undertake the task. This decision is irrevocable, and the receiver has capacity for a single task. An alternative motivation could involve incoming threats with a limited defense capacity, job candidates for a single position, priority routing for incoming jobs, and more.

Our goal is to provide good online signaling schemes that maximize the expected utility of the sender. We design simple schemes and characterize the impact of dynamic information revelation with *prophet inequalities* – we compare the performance of our online schemes to the optimal expected utility for the sender that could be obtained from knowing all the information in advance.

1.1 Our Contribution

We study online Bayesian persuasion inspired by prophet inequalities for the classic gambler's problem. At each round, $i = 1, \ldots, n$, a pair of values (one for the sender and one for the receiver) is drawn from a commonly known prior distribution, $\mathcal{D}_i$. The pair of values is revealed to the sender who then partly shares this information with the receiver according to a signaling scheme committed to in advance. Based on this information, the receiver takes an irrevocable binary decision.

We begin by considering the iid case ($\mathcal{D}_i = \mathcal{D}_j$ for all i, j) in Section 2. We show that a simple signaling scheme designed by Dughmi and Xu [2016] for the offline case can be applied online and that it provides a $(1-1/e)$-approximation. We then turn to the more general case under an additional assumption, referred to as the "satisfactory-status-quo" (or SSQ for short) assumption. Here, there exists an outside option (e.g. the current employee) for the receiver which has a deterministic value which is at least as profitable as her expected value of any of the candidates arriving online. An equivalent way of expressing this assumption is to assume that the last candidate has the maximum expected value for the receiver among all candidates. For this scenario, we provide an online signaling scheme that provides a $1/2$-approximation and show that this is the best possible guarantee. Unfortunately, if we drop the SSQ assumption, then in general no scheme can guarantee any finite approximation. On the positive side, one can compute an optimal online scheme in polynomial time.

In Section 3, we discuss extensions of our general design template. More concretely, we design online schemes for matching variants with multiple receivers and multiple candidate hires per receiver, when each receiver satisfies SSQ. Our schemes rely on an optimal solution to an LP-relaxation, combined with carefully designed probabilistic damping techniques.

1.2 Related Work

Our work extends the study of prophet inequalities which was introduced by Krengel and Sucheston [1977]. More recently, the problem was the focus of a lot of research which showed improvements for special cases and introduced combinatorial variants of the original problem [Dütting *et al.*, 2017; Kleinberg and Weinberg, 2019; Alaei, 2014; Chawla *et al.*, 2010; Correa *et al.*, 2017, 2019; Esfandiari *et al.*, 2015]. The study of Bayesian persuasion was initiated by Aumann and Maschler [1966] and came back into focus more recently following the work of Kamenica and Gentzkow [2011]. A plethora of authors worked on variants of persuasion [Celli *et al.*, 2020; Arieli and Babichenko, 2019; Ely, 2017; Ely *et al.*, 2015; Au, 2015; Dughmi and Xu, 2017], including combinations of persuasion with online learning and multi-armed bandit problems (see, e.g., Kremer *et al.* [2014]; Frazier *et al.* [2014]; Mansour *et al.* [2015] and subsequent work). For a recent survey on related algorithmic work see Dughmi [2017].

An online model of Bayesian persuasion closely related to our work was studied in our recent work [Hahn *et al.*, 2019]. We study a similar round-wise persuasion game with very different assumptions regarding the a-priori knowledge of sender and receiver. The scenario is inspired by the secretary problem – candidate values are unknown and adversarially chosen but candidates arrive in uniform random order. This uncertainty also has consequences for the definition of persuasiveness. In contrast, in this paper we assume candidate values are drawn independently from known distributions, which allows to compute expected utilities and to apply standard notions of persuasiveness. Moreover, we use significantly different techniques for analysis.

1.3 Model

In the basic version of our model, there are two players – a *sender* $\mathcal{S}$ and a *receiver* $\mathcal{R}$. There are n rounds, and in each round a candidate arrives. Candidate i has a *type* θ_i, and each type is associated with a pair of non-negative utility values, one for $\mathcal{S}$ and one for $\mathcal{R}$. The type of candidate i is drawn independently from distribution $\mathcal{D}_i$. The n distributions $\mathcal{D}_1, \ldots, \mathcal{D}_n$ are known to both players in advance and have finite support of size m. The probability that candidate i has type $\theta_i = j$ is denoted by q_{ij} for all[1] $i \in [n]$ and $j \in [m]$. The utility values of candidate i with type j for $\mathcal{R}$ and $\mathcal{S}$ are denoted by ρ_{ij} and ξ_{ij}, respectively. We use $\bar{\rho}_i = \mathbb{E}_{j \sim \mathcal{D}_i}[\rho_{ij}]$ to denote the expected utility for $\mathcal{R}$ in round i.

Upon arrival of a candidate in round i, $\mathcal{S}$ observes the type θ_i and the associated values and sends a signal σ_i to $\mathcal{R}$. The receiver knows $\mathcal{D}_1, \ldots, \mathcal{D}_n$ and signals $\sigma_1, \ldots, \sigma_{i-1}$. Upon reception of σ_i, she has to immediately decide whether she wants to hire or dismiss candidate i. The process ends once a candidate is hired or all n candidates have been dismissed. The decision made in each round is irrevocable. Each player strives to maximize its own utility of the hired candidate.

As usual in Bayesian persuasion, we assume that the sender has *commitment power* [Kamenica and Gentzkow, 2011], i.e., $\mathcal{S}$ can commit in advance to a *signaling scheme* φ. In each round i, the scheme φ takes as input the vector of observed types $(\theta_1, \ldots, \theta_i)$ and outputs a signal σ_i to $\mathcal{R}$. We restrict the set of schemes to schemes φ that are *direct* and *persuasive*: In a direct scheme, all signals are $\sigma_i \in \{\text{HIRE}, \text{NOT}\}$ for all $i \in [n]$. A persuasive scheme is one that is incentive-compatible, i.e., $\mathcal{R}$ maximizes her expected utility by following the recommendations. By a revelation-principle style argument, the assumption of a direct and persuasive scheme is without loss of generality [Kamenica and Gentzkow, 2011; Arieli and Babichenko, 2019].

We design online signaling schemes that are good for the sender, i.e., our goal is to find direct and persuasive schemes that maximize the sender's expected utility of the hired candidate. To avoid technicalities, we assume that $\mathcal{R}$ breaks ties in favor of $\mathcal{S}$. We compare the expected utility to that of an optimal (direct and persuasive) offline scheme, in which (the prophet-version of) $\mathcal{S}$ can see the full vector of realizations $(\theta_1, \ldots, \theta_n)$ in advance and signal to $\mathcal{R}$ accordingly.

In Section 2, we briefly consider the iid scenario (with $\mathcal{D}_1 = \mathcal{D}_2 = \ldots = \mathcal{D}_n$) and make the following *satisfactory-status-quo* assumption. Under such an assumption, there exists an external option $E \notin [n]$ for $\mathcal{R}$ which is always available and gives the best expected utility for $\mathcal{R}$ ($\rho_E \geq \bar{\rho}_i$ for all $i \in [n]$). An alternative formulation of this assumption is that the last candidate provides the highest expected value for $\mathcal{R}$, i.e. $\bar{\rho}_n \geq \bar{\rho}_i$ for all $i \in [n]$. Thus, $\mathcal{R}$ can always wait for the final candidate. Since we are interested in the worst case for $\mathcal{S}$, we assume that the utility of $\mathcal{S}$ for the outside option is zero.

[1] We use the short notation $[x] = \{1, 2, \ldots, x\}$ for $x \in \mathbb{N}$.

2 Single Candidate

2.1 A Simple Scheme for the IID-Case

We start by briefly discussing the case of iid distributions. Optimal persuasive signaling in the offline case, in which $\mathcal{S}$ sees all realizations $\theta_1, \ldots, \theta_n$ in advance, was studied by Dughmi and Xu [2016]. Consider an optimal offline scheme. We denote the ex-post distribution of the hired candidate by $\boldsymbol{x}^o = (x^o_{ij})_{i\in[n],j\in[m]}$. There is an optimal offline scheme that is symmetric, i.e., that yields $x^o_{ij} = x^o_{i'j} =: x^o_j$, for all rounds i, i'. Hence, the probability that a candidate in a round is recommended for hire is $\sum_j x^o_j = 1/n$. Dughmi and Xu show that the additional constraints to ensure persuasiveness can be expressed by a linear polytope, and an optimal scheme can be computed by solving a polynomial-sized LP.

In contrast to the optimal scheme, they also propose a simple approximate scheme, which relies on solving the following simpler LP with fewer constraints:

$$\begin{array}{llcll}
\text{Max.} & n\cdot\sum_j x_j\xi_j & & & \\
\text{s.t.} & n\cdot\sum_j x_j & = & 1 & \\
& x_j + (n-1)y_j & = & q_j & \forall\, j\in[m] \\
& \sum_j x_j\rho_j & \geq & \sum_j y_j\rho_j & \\
& x_j, y_j & \geq & 0 & \forall\, j\in[m]
\end{array} \quad (1)$$

The intuition is that x_j is an ex-post probability of receiving a HIRE signal on a candidate of type j in some round i, while y_j is an ex-post probability of getting a NOT signal. For this LP, the complex persuasiveness constraints are relaxed to simple ones – the first two are symmetry and consistency constraints, the third one requires that hiring a candidate upon HIRE signal is better for $\mathcal{R}$ than hiring one upon NOT signal. The objective function yields the expected utility of $\mathcal{S}$ from the hired candidate. It is easy to see that the ex-post distribution for the optimal symmetric offline scheme $\boldsymbol{x}^o$ is a feasible solution for this LP. Hence, an optimal LP solution $\boldsymbol{x}^*$ yields an upper bound on the expected utility for $\mathcal{S}$ in $\boldsymbol{x}^o$.

For every feasible solution $y_j = (q_j - x_j)/(n-1)$, where $y_j \geq 0$ whenever $x_j \leq q_j$. We plug this into the third constraint and rearrange it to $\sum_j x_j\rho_j \geq \sum_j q_j\rho_j\cdot\frac{1}{n} = \bar{\rho}\cdot\sum_j x_j$. Thus, LP (1) is equivalent to

$$\begin{array}{llcll}
\text{Max.} & n\cdot\sum_j x_j\xi_j & & & \\
\text{s.t.} & n\cdot\sum_j x_j & = & 1 & \\
& x_j & \leq & q_j & \forall\, j\in[m] \\
& \sum_j x_j\rho_j & \geq & \bar{\rho}\cdot\sum_j x_j & \\
& x_j & \geq & 0 & \forall\, j\in[m]
\end{array} \quad (2)$$

An LP-optimum $\boldsymbol{x}^*$ might not represent an ex-post distribution of some persuasive signaling scheme φ, since it adheres only to the simple, relaxed set of constraints. Dughmi and Xu propose a way to turn $\boldsymbol{x}^*$ into an persuasive signaling scheme, which we formulate as our *simple scheme*: In each round $i \in [n-1]$, observe the type θ_i. If there has been no HIRE, signal HIRE independently with probability $x^*_{\theta_i}/q_{\theta_i}$, and NOT otherwise. In round n, signal HIRE if not done so before, and NOT otherwise.

The simple scheme can be applied by $\mathcal{S}$ in the online scenario, because in round i it does not use information about future realizations. The proof of the following proposition follows from [Dughmi and Xu, 2016, Theorem 3.8].

Proposition 1. *The simple scheme is persuasive in the online setting and yields a $(1 - 1/e)$-approximation.*

2.2 Beyond IID

In case of general distributions, our first result is that an optimal online scheme can be computed in polynomial time. The approach is via backwards induction and solving $n-1$ linear programs. Interestingly, this result contrasts the conditions for an optimal offline scheme, which is known to be hard (for details see Dughmi and Xu [2016]).

Theorem 1. *An optimal persuasive signaling scheme in the online setting can be computed in polynomial time.*

Proof. The mechanism signals at most one HIRE signal. If the receiver does not hire, the mechanism never sends a HIRE signal again. As such, we can assume that the online process ends upon sending the first HIRE signal. Now consider the case that we reach round i without having sent any HIRE signal so far.

Let x_{ij} be the probability to signal HIRE upon arrival of a candidate of type j in round i. We determine the values of x_{ij} using an LP and the optimal solution for rounds $i+1, \ldots, n$. Let ξ^i and ρ^i be the expected utility for $\mathcal{S}$ and $\mathcal{R}$ from the optimal mechanism applied in rounds $i, i+1, \ldots, n$, respectively. Let $\bar{\rho}^i$ be the best expected utility for receiver in any single round $i, i+1, \ldots, n$, i.e., $\bar{\rho}^i = \max_{i'\in\{i,\ldots,n\}} \sum_j q_{i'j}\rho_{i'j}$.

In the last round it is optimal to set $x_{nj} = 1$ for all j, since it is in the interest of both $\mathcal{S}$ and $\mathcal{R}$ to hire the last candidate. Hence $\xi^n = \sum_j q_{nj}\xi_{nj}$ and $\rho^n = \sum_j q_{nj}\rho_{nj}$. Now suppose we have computed the optimal mechanism to be applied in rounds $i+1, \ldots, n$. Consider round i. The expected value of the sender is given by $q_{ij}\xi_{ij}$ if a candidate of type j arrived and she signals HIRE, or $q_{ij}\xi^{i+1}$ if a candidate of type j arrived and she signals NOT. Thus, $\mathcal{S}$ strives to maximize $\sum_j q_{ij}(x_{ij}\xi_{ij} + (1 - x_{ij})\xi^{i+1})$ or, equivalently,

$$\xi^{i+1} + \sum_j q_{ij}x_{ij}(\xi_{ij} - \xi^{i+1}) \ .$$

Clearly, this is maximized for $x_{ij} \in \{0, 1\}$ with $x_{ij} = 1$ if and only if $\xi_{ij} \geq \xi^{i+1}$. However, $\mathcal{S}$ also needs to incentivize $\mathcal{R}$ to follow the signal.

Suppose $\mathcal{R}$ gets a HIRE signal in round i. If she accepts, her conditional expectation is $\sum_j q_j x_{ij}\rho_{ij} / \sum_j q_j x_{ij}$. Upon rejection, she can only resort to the subsequent round with best expectation, i.e., $\bar{\rho}^{i+1}$. Hence, $\mathcal{R}$ accepts upon a HIRE signal if and only if

$$\sum_j q_{ij}x_{ij}(\rho_{ij} - \bar{\rho}^{i+1}) \geq 0 \ .$$

Suppose $\mathcal{R}$ gets a NOT signal in round i. If she accepts, her conditional expectation is $\sum_j q_j(1-x_{ij})\rho_{ij} \big/ \sum_j q_j(1-x_{ij})$. If she follows the mechanism, then by the inductive assumption that the mechanism is persuasive from round $i+1$ on, the best expected utility for $\mathcal{R}$ is ρ^{i+1}. Hence, $\mathcal{R}$ rejects upon a NOT signal if and only if $\left(\sum_j q_{ij}(1-x_{ij})\right) \cdot \rho^{i+1} \geq \sum_j q_{ij}(1-x_{ij})\rho_{ij}$, which is equivalent to

$$\sum_j q_{ij}x_{ij}(\rho_{ij} - \rho^{i+1}) \geq \sum_j q_{ij}\rho_{ij} - \rho^{i+1} \quad .$$

Given the optimal mechanism for rounds $i+1,\ldots,n$ we obtain the optimal mechanism for rounds $i,\ldots,n$ by solving the LP

$$\begin{array}{lrcl} \text{Max.} & \xi^{i+1} + \sum_j q_{ij}x_{ij}(\xi_{ij} - \xi^{i+1}) & & \\ \text{s.t.} & \sum_j q_{ij}x_{ij}(\rho_{ij} - \bar{\rho}^{i+1}) & \geq & 0 \\ & \sum_j q_{ij}x_{ij}(\rho_{ij} - \rho^{i+1}) & \geq & \sum_j q_{ij}\rho_{ij} - \rho^{i+1} \\ & x_{ij} & \in & [0,1] \;\; \forall\, j \in [m] \;\; . \end{array} \tag{3}$$

By the inductive assumption, the mechanism is persuasive for rounds $i+1,\ldots,n$. This implies $\rho^{i+1} \geq \bar{\rho}^{i+1}$, since following the mechanism must be at least as profitable for $\mathcal{R}$ as deviating to pick the candidate in the remaining round with highest expectation. Given this property, we observe that LP (3) always has a feasible solution. If $\sum_j q_{ij}\rho_{ij} \geq \bar{\rho}^{i+1}$, then $x_{ij} = 1$ for all $j \in [m]$ satisfies both constraints. Otherwise, if $\sum_j q_{ij}\rho_{ij} < \bar{\rho}^{i+1}$, then $0 > \sum_j q_{ij}\rho_{ij} - \bar{\rho}^{i+1} \geq \sum_j q_{ij}\rho_{ij} - \rho^{i+1}$, and setting $x_{ij} = 0$ for all $j \in [m]$ satisfies both constraints. □

Although we can compute the optimal online mechanism in polynomial time, there are instances in which no finite approximation to the optimal offline signaling scheme can be obtained. This is a consequence of commitment power of the sender – $\mathcal{S}$ publishes and commits to a signaling scheme φ in advance. By inspecting this scheme, $\mathcal{R}$ can determine if $\mathcal{S}$ in round i uses access to $(\theta_1,\ldots,\theta_n)$ available in the offline case or $(\theta_1,\ldots,\theta_i)$ available in the online case. Hence, $\mathcal{R}$ can determine whether the signal of $\mathcal{S}$ contains information about realizations in future rounds or not. This property is key for the following lower bound.

Theorem 2. *There are instances in which the optimal online signaling scheme yields an unbounded approximation ratio.*

Proof. Consider the following instance with $n = 2$ candidates. $\mathcal{D}_1$ is deterministic, a single realization with value pair $(\rho_{11}, \xi_{11}) = (1, 0)$. $\mathcal{D}_2$ has two possible realizations with probability $1/2$ each. The value pairs are $(\rho_{21}, \xi_{21}) = (2-\varepsilon, 1)$ for some $\varepsilon \in (0,1)$, and $(\rho_{22}, \xi_{22}) = (0,0)$.

Consider the optimal offline scheme, in which the sender knows both realizations. $\mathcal{S}$ signals HIRE for the second candidate if and only if the realization $\theta_2 = (2-\varepsilon, 1)$. Otherwise, $\mathcal{S}$ signals HIRE for the first candidate. In this way, $\mathcal{R}$ always gets her optimal candidate – the scheme is persuasive. The expected utility for $\mathcal{S}$ is $1/2$.

Now consider the online case. By inspecting the online scheme, the receiver realizes that the signal in the first round contains no information – θ_1 is perfectly known to both $\mathcal{S}$ and $\mathcal{R}$, and $\mathcal{S}$ has no information about θ_2. The signal in the second round, however, is irrelevant for $\mathcal{R}$ – upon reaching round 2, it is a dominant strategy for $\mathcal{R}$ to hire the second candidate. As a consequence, $\mathcal{R}$ will take an action independent of the signal of $\mathcal{S}$ and accept the candidate in round 1, since it yields the higher expected value. The unique persuasive scheme in the online scenario for $\mathcal{S}$ is to signal HIRE in the first round, which has utility 0 for $\mathcal{S}$. □

There is a broad set of conditions under which the online case leads to a drastic deterioration in expected sender utility. For example, if a later candidate has a "golden-nugget"-type distribution (small expected value, with tiny probability a super-valuable realization for both players), then an offline scheme can convince $\mathcal{R}$ to wait, since the signal contains the information that the "golden nugget" will indeed arrive. In contrast, an online scheme cannot transport this information, and hence $\mathcal{R}$ has an incentive to accept an earlier candidate with better expected value for her (and possibly much less value for $\mathcal{S}$).

2.3 A Simple Scheme for SSQ

In this section, we extend the idea of the simple scheme in the case that SSQ holds. The main condition to ensure a small constant approximation ratio is that the receiver has a canonical option for deviation. In the iid case a valid deviation from a HIRE signal in rounds $i \in [n-1]$ is to simply take the last candidate in round n. This candidate has the best (in fact, the same) expected value for $\mathcal{R}$ as every other candidate $i+1,\ldots,n-1$. In this section, we assume that there is an external option $E \notin [n]$ that has the best expectation for $\mathcal{R}$, i.e., $\rho_E \geq \bar{\rho}_i$ for all $i = 1,\ldots,n$. This external option (i.e. the current employee to be replaced) can be chosen by $\mathcal{R}$ at any time. Equivalently, we assume that the last candidate has the highest expectation for $\mathcal{R}$, i.e. $\bar{\rho}_n \geq \bar{\rho}_i$, for all rounds $i \in [n-1]$.

An optimal offline scheme can be obtained by solving an exponential-sized LP using the ideas in [Dughmi and Xu, 2016]. Instead, we set up a polynomial-sized LP as a natural extension of LP (2). The intuition for x_{ij} is the ex-post probability for a HIRE signal for candidate type j from $\mathcal{D}_i$:

$$\begin{array}{lrcll} \text{Max.} & \sum_{i,j} x_{ij}\xi_{ij} & & & \\ \text{s.t.} & \sum_{i,j} x_{ij} & \leq & 1 & \\ & x_{ij} & \leq & q_{ij} & \forall\, i \in [n], j \in [m] \\ & \sum_j x_{ij}\rho_{ij} & \geq & \rho_E \cdot \sum_j x_{ij} & \forall\, i \in [n] \\ & x_{ij} & \geq & 0 & \forall\, i \in [n], j \in [m] \end{array} \tag{4}$$

Lemma 1. *The optimal value of LP* (4) *is an upper bound on the expected utility for $\mathcal{S}$ in any offline persuasive signaling scheme.*

Proof. Consider an optimal offline scheme for $\mathcal{S}$. Suppose x_{ij} be the ex-post probability that a HIRE signal is issued for candidate j from $\mathcal{D}_i$. We show that the vector $\boldsymbol{x}$ defined in this way is a feasible solution for LP (4). The utility function of the LP corresponds to the expected utility for $\mathcal{S}$ upon a successful recommendation. The constraint $x_{ij} \leq q_{ij}$ is fulfilled, since a candidate cannot be recommended more often than it arrives. The constraint $\sum_j x_{ij}\rho_{ij} \geq \rho_E \cdot \sum_j x_{ij}$ is fulfilled for all $i \in [n]$, since the scheme is persuasive and therefore does not allow a profitable deviation upon a HIRE signal in round i to the outside option. The remaining two constraints simply state that $\boldsymbol{x}$ is a vector of probabilities that sum to at most 1 – both conditions are fulfilled in the optimal offline scheme. Thus, $\boldsymbol{x}$ is a feasible solution for the LP. The optimal solution to LP (4) gives an upper bound on the expected utility for $\mathcal{S}$ in any persuasive scheme. □

Based on the optimal solution $\boldsymbol{x}^*$ to LP (4), we define a *simple scheme for SSQ* (Algorithm 1) and prove our main result of this section.

Algorithm 1: Simple Scheme for SSQ

Input: Distributions $(\mathcal{D}_i)_{i\in[n]}$, factors $\boldsymbol{d} = (d_i)_{i\in[n]}$, online sequence of θ_i drawn from $\mathcal{D}_i$
for *rounds* $i = 1$ *to* n **do**
  Upon seeing the draw θ_i from $\mathcal{D}_i$:
  W. prob. $1 - d_i$: Signal NOT, go to next round.
  Otherwise, w. prob. $x^*_{i\theta_i}/q_{i\theta_i}$: Signal HIRE now and NOT in all remaining rounds
  Otherwise: Signal NOT, go to next round

Theorem 3. *For a suitable choice of parameters* **d***, the simple scheme is persuasive in the online setting satisfying SSQ and yields a* $1/2$*-approximation.*

Proof. First, we show that for every set of damping parameters $\boldsymbol{d} \in [0,1]^n$ the scheme is persuasive. We subdivide the proof into two cases. In case 1 we assume $\mathcal{S}$ signals HIRE in round i and NOT in all previous rounds. By following the signal, Bayes' rule shows that $\mathcal{R}$ gets an expected utility of

$$\frac{\sum_j q_{ij} \cdot x_{ij} \cdot d_i \cdot \rho_{ij}/q_{ij}}{\sum_j q_{ij} \cdot x_{ij} \cdot d_i/q_{ij}} = \frac{\sum_j x_{ij}\rho_{ij}}{\sum_j x_{ij}} \geq \rho_E \ ,$$

where the last inequality follows from the LP-constraint $\sum_j x_{ij}\rho_{ij} \geq \rho_E \cdot \sum_j x_{ij}$. In case the receiver decides to deviate and reject candidate θ_i, the scheme will not provide any more signals. Thus, ρ_E is the maximum utility that $\mathcal{R}$ can expect for such a deviation. Hence, $\mathcal{R}$ has an incentive to follow the signal.

In case 2, we assume that there are only signals NOT in all rounds $1,\ldots,i$. If the receiver deviates and hires, she gets an expected utility of

$$\frac{\sum_j q_{ij} \cdot \left(1 - \frac{x_{ij}d_i}{q_{ij}}\right) \cdot \rho_{ij}}{\sum_j q_{ij}\left(1 - \frac{x_{ij}d_i}{q_{ij}}\right)} = \frac{\sum_j (q_{ij} - x_{ij}d_i) \cdot \rho_{ij}}{\sum_j q_{ij} - x_{ij}d_i}$$

$$= \frac{\sum_j q_{ij}\rho_{ij} - d_i \cdot \sum_j x_{ij}\rho_{ij}}{1 - d_i \sum_j x_{ij}} \leq \frac{\rho_E - d_i\rho_E \sum_j x_{ij}}{1 - d_i \sum_j x_{ij}}$$

$$= \rho_E \ .$$

The inequality arises from the SSQ-constraint $\sum_j q_{ij}\rho_{ij} = \bar{\rho}_i \leq \rho_E$ and the LP-constraint $\sum_j x_{ij}\rho_{ij} \geq \rho_E \cdot \sum_j x_{ij}$. If, on the other hand, $\mathcal{R}$ obeys the signal and does not hire, the expected utility from the remaining scheme in rounds $i+1,\ldots,n$ is at least ρ_E: If $\mathcal{R}$ gets a HIRE signal in one of the subsequent rounds $i' > i$, we showed in case 1 that the conditional expectation upon hiring in round i' is at least ρ_E. Otherwise, if $\mathcal{R}$ gets no HIRE signal in the later rounds, she can revert to the outside option and secure a value of ρ_E. This shows that the scheme is persuasive.

Let us now show that it achieves a $1/2$-approximation w.r.t. the optimal offline signaling scheme. Using Lemma 1, it is sufficient to show that the scheme achieves an expected utility for $\mathcal{S}$ that is at least $1/2$ of the optimal value for LP (4).

Following Chawla *et al.* [2010] as well as Alaei [2014], we use damping factors d_i for all $i \in [n]$ defined as follows. We define $r_i = \Pr[\text{reaching round } i]$. It follows a recursion: $r_1 = 1, r_{i+1} = r_i\left(1 - d_i \sum_j x_{ij}\right)$. We choose $d_i = 1/(2r_i)$, which yields $r_i \cdot d_i = 1/2$ for every $i \in [n]$. d_i is well-defined since, inductively, $r_i \geq 1/2$. This is due to $r_{i+1} = r_1 - \frac{1}{2} \cdot \sum_{k\leq i,j} x_{kj}$ and $\sum_{i,j} x_{ij} \leq 1$. As a consequence, $\mathcal{S}$ obtains an expected utility of

$$\sum_i r_i \cdot d_i \cdot \sum_j \xi_{ij} \cdot q_{ij} \cdot x^*_{ij}/q_{ij}$$

$$= \sum_i r_i \cdot d_i \cdot \sum_j x^*_{ij}\xi_{ij} = \frac{1}{2}\sum_{i,j} x^*_{ij}\xi_{ij} \ ,$$

i.e., $1/2$ of the optimal value of LP (4). □

The bound of 1/2 is best possible: Suppose $n = 2$, $\mathcal{D}_1$ is deterministic with $(\rho_{11}, \xi_{11}) = (1,1)$, $\mathcal{D}_2$ has two realizations $((\rho_{21}, \xi_{21}) = (n,n)$ with probability $1/n$ and $(\rho_{22}, \xi_{22}) = (0,0)$ with probability $1 - 1/n)$, and $\rho_E = 1$. The optimal offline scheme recommends the best candidate. It yields expected utility of $2 - 1/n$ for both $\mathcal{S}$ and $\mathcal{R}$. Online schemes can only guarantee a utility of at most 1.

3 Extensions

The approach in the previous section can be generalized to a variety of combinatorial problems when a good external option provides a canonical deviation opportunity for $\mathcal{R}$. In this case, to obtain persuasiveness it suffices to provide $\mathcal{R}$ with an expected value of ρ_E conditioned on a HIRE signal. We discuss natural extensions with multiple hirings and multiple receivers. Our approach is again to solve an appropriate LP-relaxation for the ex-post distribution of hired candidates and compute a signal using carefully chosen probabilistic damping factors. For the latter we incorporate results from the area of Bayesian mechanism design with sequential posted prices [Alaei, 2014].

3.1 Hiring Multiple Candidates

Suppose $\mathcal{R}$ strives to hire $1 \le k \le n$ candidates and has k good outside options, i.e., k options with a value ρ_E each, where $\rho_E \ge \bar{\rho}_i$ for all $i \in [n]$. We call this k-*SSQ* for short. The utilities of both $\mathcal{S}$ and $\mathcal{R}$ are additive over the hired candidates.

Theorem 4. *There is a persuasive signaling scheme in the online setting with k hires satisfying k-SSQ that yields a $\left(1 - \frac{1}{\sqrt{k+3}}\right)$-approximation.*

Proof. Consider LP (5) as the natural extension of LP (4). The optimal value constitutes an upper bound for the expected utility for $\mathcal{S}$ in the optimal offline scheme – if we set x_{ij} to the ex-post probability of hiring candidate i of type j in the offline scheme, the vector $\boldsymbol{x}$ is feasible for the LP and the objective function value is the expected utility of $\mathcal{S}$ in the offline scheme.

$$\begin{array}{llcll} \text{Max.} & \sum_{i,j} x_{ij}\xi_{ij} & & & \\ \text{s.t.} & \sum_{i,j} x_{ij} & \le & k & \\ & x_{ij} & \le & q_{ij} & \forall\, i \in [n], j \in [m] \\ & \sum_{j} x_{ij}\rho_{ij} & \ge & \rho_E \sum_{j} x_{ij} & \forall\, i \in [n] \\ & x_{ij} & \ge & 0 & \forall\, i \in [n], j \in [m] \end{array} \tag{5}$$

To devise an online scheme, we use a natural extension of the simple scheme under SSQ. First solve LP (5) optimally, let $\boldsymbol{x}^*$ be the optimum solution. The decision in round i is again split into two steps. In step 1, send NOT with probability $1 - d_i$, otherwise advance to step 2. In step 2, send HIRE with probability $x^*_{i\theta_i}/q_{i\theta_i}$, and NOT otherwise. Overall, we send HIRE with probability $d_i \cdot x^*_{i\theta_i}/q_{i\theta_i}$.

The exact same calculations as in Theorem 3 show that conditioned on a HIRE signal in round i, the expected value of the candidate for $\mathcal{R}$ is at least ρ_E. The same calculations show that conditioned on a NOT signal in round i, the expected value of the candidate for $\mathcal{R}$ is at most ρ_E. Hence, a deviation of $\mathcal{R}$ to an outside option or to hiring upon a NOT signal is not profitable and, thus, the scheme is persuasive.

To show the approximation factor, we need to carefully design the damping scheme d_i for step 1. d_i should be high to hire good candidates in round i, but also low to ensure that better candidates in later rounds $i' > i$ can be hired. This is exactly the trade-off faced by the "γ-Conservative Magician" in [Alaei, 2014]. Here, a magician has k wands to open n boxes. If he applies a wand to a box, the box opens, but the wand breaks with some probability. Each box comes with a distribution for the value of the content and a probability that a wand breaks when opening it. In our context, the k wands are k positions for hire, the n boxes are n rounds, "opening a box" means surviving step 1, and "breaking the wand" means actually sending a HIRE signal in step 2.

In [Alaei, 2014, Definition 3] Alaei devises an adaptive strategy to set the d_i such that every round i the joint probability of having at most $k-1$ HIRE signals and arriving at step 2 of round i is at least $\gamma_k = 1 - \frac{1}{\sqrt{k+3}}$. Hence, using his strategy to design the d_i, the expected utility for the sender in the resulting scheme is at least

$$\sum_i \gamma_k \sum_j \xi_{ij} \cdot q_{ij} \cdot x^*_{ij}/q_{ij} = \gamma_k \sum_{i,j} x^*_{ij}\xi_{ij} \;.$$

It recovers at least a γ_k-fraction of the optimum for LP (5). □

3.2 Hiring with Multiple Receivers

Private Signals Suppose there are ℓ different receivers $\mathcal{R}_1, \dots, \mathcal{R}_\ell$. In every round i and for every $t \in [\ell]$, $\mathcal{S}$ sends $\mathcal{R}_t$ a private signal $\sigma_i^{(t)} \in \{\text{HIRE}, \text{NOT}\}$ (c.f. [Dughmi and Xu, 2017] for results on private and public signaling channels in a related offline scenario). A problem arising with multiple receivers is feasibility of the assignment of candidates. There might be conflicting situations when several receivers simultaneously decide to hire the same candidate. We circumvent this problem due to the following: (1) Our scheme will ensure at most a single receiver gets a HIRE-signal in every round. (2) When deviating, the (weakly) most-preferred option for any receiver is the outside option, which does not interfere with the other receivers. Note that our scheme is persuasive even when every receiver assumes that she can hire every candidate, irrespective of the decisions made by others.

Formally, a candidate type is a vector $(\rho_{ijt}, \xi_{ijt})_{t \in [\ell]}$ of value-pairs, for all $i \in [n], j \in [m]$. Receiver $\mathcal{R}_t$ strives to hire k_t candidates and fulfills k_t-SSQ, i.e. has k_t unique good outside options with value $\rho_E^{(t)} \ge \bar{\rho}_i^{(t)}$, where $\bar{\rho}_i^{(t)} = \mathbb{E}_{j \sim \mathcal{D}_i}[\rho_{ijt}]$. We define $k = \min_t k_t$. The utility of the sender is additive over all hired candidates. The utility of receiver $\mathcal{R}_t$ is additive over candidates hired by $\mathcal{R}_t$.

Theorem 5. *There is a persuasive signaling scheme in the online setting with ℓ receivers $\mathcal{R}_1, \dots, \mathcal{R}_\ell$ with $\mathcal{R}_t$ having k_t hires and satisfying k_t-SSQ for all $t \in [\ell]$ that yields a $\left(1 - \frac{1}{\sqrt{k+3}}\right)$-approximation.*

Proof. LP (6) is an extension of LP (5), where the optimal value again constitutes an upper bound for the expected utility for $\mathcal{S}$ in the optimal offline scheme – the ex-post probability of $\mathcal{R}_t$ hiring candidate i of type j in the offline scheme yields a feasible solution.

$$\begin{array}{llcll} \text{Max.} & \sum_{i,j,t} x_{ijt}\xi_{ijt} & & & \\ \text{s.t.} & \sum_{i,j} x_{ijt} & \le & k_t & \forall\, t \in [\ell] \\ & \sum_{t} x_{ijt} & \le & q_{ij} & \forall\, i \in [n], j \in [m] \\ & \sum_{j} x_{ijt}\rho_{ijt} & \ge & \rho_E^{(t)} \sum_{j} x_{ijt} & \forall\, i \in [n], t \in [\ell] \\ & x_{ijt} & \ge & 0 & \forall\, i, j, t \end{array} \tag{6}$$

For the online scheme, we extend the simple schemes studied above. First solve LP (6) optimally, let $\boldsymbol{x}^*$ be the optimum solution. In decision step 1, we independently send NOT to $\mathcal{R}_t$ with probability $1 - d_i^{(t)}$ for all $t \in [\ell]$. In decision step 2, we make a correlated choice and pick at most one receiver for a HIRE signal as follows: With probability $x^*_{i\theta_i t}/q_{i\theta_i}$, choose $\mathcal{R}_t$. With probability $1 - \sum_t x^*_{i\theta_i t}/q_{i\theta_i}$, do not select any receiver. Iff $\mathcal{R}_t$ is drawn in step 2 and has not received a NOT-signal in step 1, we signal HIRE to $\mathcal{R}_t$. This ensures that every candidate is recommended for hire to at most one receiver. Hence, overall we send HIRE for candidate i of type j with probability $d_i \cdot \sum_t x^*_{i\theta_i t}/q_{i\theta_i}$.

The exact same calculations as in Theorem 4 show that conditioned on a HIRE signal for $\mathcal{R}_t$ in round i, the expected value of the candidate for $\mathcal{R}_t$ is at least $\rho_E^{(t)}$. The same calculations show that conditioned on a NOT signal for $\mathcal{R}_t$ in round i, the expected value of the candidate for $\mathcal{R}_t$ is at most $\rho_E^{(t)}$. Hence, the scheme is persuasive for all receivers $\mathcal{R}_1, \ldots, \mathcal{R}_\ell$. A deviation to the outside option is always (weakly) preferable for any receiver. Hence, we can assume that the sets of candidates hired by different receivers are distinct.

To show the approximation factor, we observe that the proof approach presented for Theorem 4 applies. The different receivers represent different magicians, each having its own supply of wands. Only if a HIRE-signal is sent to $\mathcal{R}_t$ in step 2, one of the wands of this magician breaks. This ensures that each candidate will only be hired by at most one receiver. The correlated choice of sending a HIRE signal (i.e., breaking a wand of some receiver) does not change the argument.

Now since $\gamma_k = \left(1 - \frac{1}{\sqrt{k+3}}\right)$ is non-decreasing in k, the expected utility for $\mathcal{S}$ is

$$\sum_t \gamma_{k_t} \sum_{i,j} \xi_{ijt} \cdot q_{ij} \cdot x^*_{ijt}/q_{ij} \geq \gamma_k \sum_{i,j,t} x^*_{ijt} \xi_{ijt} \quad ,$$

i.e., at least a γ_k-fraction of the optimum of LP (6).

□

Public Signals Let us briefly look at the case when $\mathcal{S}$ can only use *public signals* (instead of private signals studied above). With public signals $\mathcal{S}$ faces the feasibility problem discussed above, i.e., multiple receivers might get an incentive to hire the same candidate. Ensuring feasibility with persuasive signals can lead to a drastic performance loss for $\mathcal{S}$, even in case of SSQ for all receivers.

Proposition 2. *There are instances with SSQ in which any persuasive signaling scheme with public signals yields an unbounded approximation ratio.*

Proof. Consider 2 candidates and 2 receivers as follows. Both candidates have a high and a low type, each occurring with probability $1/2$. The low type has values $(\rho_{i21}, \xi_{i21}) = (\rho_{i22}, \xi_{i22}) = (0, 0)$ for both candidates $i = 1, 2$. The first candidate's high type has values $(\rho_{111}, \xi_{111}) = (1, M)$ and $(\rho_{112}, \xi_{112}) = (3/4 + \varepsilon, 0)$, where $M > 1$ is a large number. The second candidate's high type has values $(\rho_{211}, \xi_{211}) = (0, 0)$ and $(\rho_{212}, \xi_{212}) = (1, 1)$. The outside options are $\rho_E^{(1)} = 1$, $\rho_E^{(2)} = 3/4$.

With probability $1/4$, both candidates have the high type. Hence, in the offline scenario, the sender can signal HIRE for candidate 1 to $\mathcal{R}_1$ and candidate 2 to $\mathcal{R}_2$. This is persuasive and yields a utility of $M + 1$. $\mathcal{S}$ sends NOT signals to both receivers in all other cases. Overall, the expected utility is at least $(M + 1)/4$.

It is crucial for $\mathcal{S}$ that it becomes sufficiently likely that $\mathcal{R}_1$ hires candidate 1 when it has the high type. In the online scenario, however, the only possibility to motivate $\mathcal{R}_1$ to do so is when the HIRE signal for candidate 1 implies it has high type with certainty. Since the signal is public, $\mathcal{R}_2$ also knows with certainty that candidate 1 has high type and, hence, also has an incentive to hire. Therefore, in any persuasive online scheme $\mathcal{S}$ can never motivate $\mathcal{R}_1$ (alone) to hire candidate 1 when it has the high type. □

References

Saeed Alaei. Bayesian combinatorial auctions: Expanding single buyer mechanisms to many buyers. *SIAM J. Comput.*, 43(2):930–972, 2014.

Ricardo Alonso and Odilon Câmara. Persuading voters. *Amer. Econ. Rev.*, 106(11):3590–3605, 2016.

Itai Arieli and Yakov Babichenko. Private bayesian persuasion. *J. Econ. Theory*, 182:185–217, 2019.

Pak Hung Au. Dynamic information disclosure. *RAND J. Econ.*, 46(4):791–823, 2015.

Robert Aumann and Michael Maschler. Game theoretic aspects of gradual disarmament. *Report of the US Arms Control and Disarmament Agency*, 80:1–55, 1966.

Ashwinkumar Badanidiyuru, Kshipra Bhawalkar, and Haifeng Xu. Targeting and signaling in ad auctions. In *Proc. 29th Symp. Disc. Algorithms (SODA)*, pages 2545–2563, 2018.

Andrea Celli, Stefano Coniglio, and Nicola Gatti. Private Bayesian persuasion with sequential games. In *Proc. 34th Conf. Artif. Intell. (AAAI)*, 2020.

Shuchi Chawla, Jason Hartline, David Malec, and Balasubramanian Sivan. Multi-parameter mechanism design and sequential posted pricing. In *Proc. 42nd Symp. Theory Comput. (STOC)*, pages 311–320, 2010.

José Correa, Patricio Foncea, Ruben Hoeksma, Tim Oosterwijk, and Tjark Vredeveld. Posted price mechanisms for a random stream of customers. In *Proc. 18th Conf. Econ. Comput. (EC)*, pages 169–186, 2017.

José Correa, Paul Dütting, Felix Fischer, and Kevin Schewior. Prophet inequalities for I.I.D. random variables from an unknown distribution. In *Proc. 20th Conf. Econ. Comput. (EC)*, pages 3–17, 2019.

Shaddin Dughmi and Haifeng Xu. Algorithmic Bayesian persuasion. In *Proc. 48th Symp. Theory Comput. (STOC)*, pages 412–425, 2016.

Shaddin Dughmi and Haifeng Xu. Algorithmic persuasion with no externalities. In *Proc. 18th Conf. Econ. Comput. (EC)*, pages 351–368, 2017.

Shaddin Dughmi. Algorithmic information structure design: a survey. *SIGecom Exchanges*, 15(2):2–24, 2017.

Paul Dütting, Michal Feldman, Thomas Kesselheim, and Brendan Lucier. Prophet inequalities made easy: Stochastic optimization by pricing non-stochastic inputs. In *Proc. 27th Symp. Found. Comput. Sci. (FOCS)*, pages 540–551, 2017.

Jeffrey Ely, Alexander Frankel, and Emir Kamenica. Suspense and surprise. *J. Political Econ.*, 123(1):215–260, 2015.

Jeffrey Ely. Beeps. *Amer. Econ. Rev.*, 107(1):31–53, 2017.

Yuval Emek, Michal Feldman, Iftah Gamzu, Renato Paes Leme, and Moshe Tennenholtz. Signaling schemes for revenue maximization. In *Proc. 13th Conf. Econ. Comput. (EC)*, pages 514–531, 2012.

Hossein Esfandiari, MohammadTaghi Hajiaghayi, Vahid Liaghat, and Morteza Monemizadeh. Prophet secretary. In *Proc. 23rd European Symp. Algorithms (ESA)*, pages 496–508, 2015.

Peter Frazier, David Kempe, Jon Kleinberg, and Robert Kleinberg. Incentivizing exploration. In *Proc. 15th Conf. Econ. Comput. (EC)*, pages 5–22, 2014.

Itay Goldstein and Yaron Leitner. Stress tests and information disclosure. *J. Econ. Theory*, 177:34–69, 2018.

Niklas Hahn, Martin Hoefer, and Rann Smorodinsky. The secretary recommendation problem. *CoRR*, abs/1907.04252, 2019.

Emir Kamenica and Matthew Gentzkow. Bayesian persuasion. *Amer. Econ. Rev.*, 101(6):2590–2615, 2011.

Robert Kleinberg and Matthew Weinberg. Matroid prophet inequalities and applications to multi-dimensional mechanism design. *Games Econom. Behav.*, 113:97–115, 2019.

Anton Kolotilin. Experimental design to persuade. *Games Econom. Behav.*, 90:215 – 226, 2015.

Ilan Kremer, Yishay Mansour, and Motty Perry. Implementing the wisdom of the crowd. *J. Political Econ.*, 122:988–1012, 2014.

Ulrich Krengel and Louis Sucheston. Semiamarts and finite values. *Bull. Amer. Math. Soc*, 83:745–747, 1977.

Yishay Mansour, Aleksandrs Slivkins, and Vasilis Syrgkanis. Bayesian incentive-compatible bandit exploration. In *Proc. 16th Conf. Econ. Comput. (EC)*, pages 565–582, 2015.

Zinovi Rabinovich, Albert Xin Jiang, Manish Jain, and Haifeng Xu. Information disclosure as a means to security. In *Proc. 14th Conf. Auton. Agents and Multi-Agent Syst. (AAMAS)*, pages 645–653, 2015.

Haifeng Xu, Zinovi Rabinovich, Shaddin Dughmi, and Milind Tambe. Exploring information asymmetry in two-stage security games. In *Proc. 29th Conf. Artif. Intell. (AAAI)*, pages 1057–1063, 2015.

Haifeng Xu, Rupert Freeman, Vincent Conitzer, Shaddin Dughmi, and Milind Tambe. Signaling in Bayesian Stackelberg games. In *Proc. 15th Conf. Auton. Agents and Multi-Agent Syst. (AAMAS)*, pages 150–158, 2016.