Optimal Prophet Inequality with Less than One Sample

Nick Gravin* Hao Li* Zhihao Gavin Tang*

Abstract

There is a growing interest in studying sample-based prophet inequality with the motivation stemming from the connection between the prophet inequalities and the sequential posted pricing mechanisms. Rubinstein, Wang, and Weinberg (ITCS 2021) established the optimal single-choice prophet inequality with only a single sample per each distribution. Our work considers the sample-based prophet inequality with *less than one sample* per distribution, i.e., scenarios with no prior information about some of the random variables. Specifically, we propose a p-sample model, where a sample from each distribution is revealed with probability $p \in [0, 1]$ independently across all distributions. This model generalizes the single-sample setting of Rubinstein, Wang, and Weinberg (ITCS 2021), and the i.i.d. prophet inequality with a linear number of samples of Correa et al. (EC 2019). Our main result is the optimal $\frac{p}{1+p}$ prophet inequality for all $p \in [0, 1]$.

1 Introduction

Prophet inequality is a fundamental problem in optimal stopping theory and online Bayesian optimization. Consider a sequence of n boxes arriving online, each box i associated with a random value X_i sampled from a priori known distribution $\mathcal{D}_i$. The actual value of X_i is observed upon the arrival of the box i and the algorithm decides immediately whether to accept it. If the box is accepted, the algorithm collects the observed value X_i and the game ends. Else, the algorithm proceeds to the next box. The goal is to maximize the value of the accepted box and to compete against the expected maximum value of all boxes, i.e., $\mathbb{E}[\max_i X_i]$. The benchmark is also known as the prophet, since it can be interpreted as the expected value of an optimal algorithm that can look into the values of all boxes before making a choice. Krengel and Sucheston [23, 24] first established an optimal $\frac{1}{2}$-competitive prophet inequality. Subsequently, Samuel-Cahn [27] provided a single-threshold algorithm with the same tight competitive ratio.

The classic single-choice prophet inequality is equivalent to the problem of designing revenue-maximizing sequential posted pricing mechanism [20, 10]. That connection has inspired a number of generalizations in the field of algorithmic mechanism design to multi-choice settings such as matroids [4, 22, 16], matchings and combinatorial auctions [15, 18, 14], and general downward-closed constraints [25]. Furthermore, the sequential posted pricing motivates the study of prophet inequalities with limited information, as the complete knowledge of the prior distributions $(\mathcal{D}_i)_{i=1}^n$ is a rather strong and unrealistic assumption as was pointed out by Azar, Kleinberg, and Weinberg [1].

In the limited information setting, the algorithm may only access a limited number of samples per each distribution $\mathcal{D}_i$ instead of the complete description of $\mathcal{D}_i$ in the full-information case. This model is arguably more realistic than the full-information model, since samples are easy to collect, e.g., from historical data. Azar, Kleinberg, and Weinberg designed constant competitive algorithms

*ITCS, Shanghai University of Finance and Economics, Email: {nikolai, tang.zhihao}@mail.shufe.edu.cn, sufelihao@outlook.com

with only a constant number of samples per distribution for various matroid and matching settings. Recently, Rubinstein, Wang, and Weinberg [26] proved that the optimal 1/2 competitive ratio for the single-choice prophet inequality can be achieved, with only a *single* sample per distribution. Furthermore, Caramanis et al. [3] explored the limit of single-sample prophet inequalities for matroids, matching, and combinatorial auctions. Correa et al. [5] also studied the prophet secretary problem with a single sample per distribution and obtained a 0.635-competitive algorithm.

1.1 Our Contributions

Model and Result. In the regime with little prior information, it is reasonable to assume that some distributions may be completely new, i.e., they have no samples whatsoever. We propose a new framework of p-sparse sample access parameterized by $p \in [0,1]$ and apply it to the single-choice prophet inequality. Specifically, we assume that independently for each box, the algorithm sees a sample from it with probability p. It generalizes the single-sample setting of Rubinstein, Wang, and Weinberg [26] when $p = 1$.

Our *less than one sample* regime also generalizes the model of Correa et al. [7] who studied the setting with linear βn and sublinear $o(n)$ number of samples for n *i.i.d. distributions*. They showed that no algorithm can achieve a competitive ratio better than $1/e$ when $\beta = o(1)$, and designed a $(1-1/e)$-competitive algorithm for $\beta = 1$. Subsequently, Correa et al. [8] achieved a tight competitive ratio of $\frac{1+\beta}{e}$ for all $\beta \leq \frac{1}{e-1}$, and improved the competitive ratio to 0.648 for $\beta = 1$.

To the best of our knowledge, we are the first to study the *less than one sample* setting for non identical distributions. Our main result is a *tight* $\frac{p}{1+p}$-competitive algorithm for the single-choice prophet inequality. Our algorithm is based on the simple Maximum-sample algorithm [26] that stops at the first value box i with a value X_i greater than the maximum sample. However, our version (see Definition 2.1 later in the paper) has a non trivial alternation.

Techniques. First, note that stochastic optimization with a constant number of samples makes problem so much harder than the full information case, e.g., in the closely related auction literature on revenue maximization with samples, designing mechanisms with 1 sample per distribution [12, 17] is quite different from the full information setting. Moreover, it is a daunting task to get an improvement on the revenue guarantee from the setting with 1 sample to 2 samples per distribution (see, e.g., [2, 11]).

At the technical level our analysis proceeds by reducing the original problem of maximizing the expected value to a simpler objective of stopping at any of the top k card values for each fixed $k \in \mathbb{N}$. We first studied the problem for $k = 1$ and identified a hard family of instances that already gives a tight upper bound of $\frac{p}{1+p}$ on the competitive ratio for any $p \in (0,1]$.

Next, we found the right variation of the Max-Sample algorithm for the objective of stopping at the maximum ($k = 1$) with a matching lower bound of $\frac{p}{1+p}$ on the competitive ratio. Our proof proceeds by carefully constructing a set of disjoint events that would guarantee Max-Sample to win. A challenging part was to define/select the events in such a way that would keep the number of cases at a minimum. Lastly, we extended the analysis for $k = 1$ to arbitrary $k \in \mathbb{N}$ with a noticeably more elaborate set of winning events and larger case analysis.

1.2 Further Related Works

A closely related problem to prophet inequality is the celebrated secretary problem. In this setting, n elements arrive in a random arrival order. An online algorithm observes the value of each element and decide whether to take it immediately and irrevocably. Observe that in this setting, the algorithm has no prior information of the n values. Recently, Kaplan, Naori, and Raz [21] proposed a data-driven variant to the secretary problem. They assume that among the n values, a fraction p of the values are given as samples to the algorithm in advance; the remaining values either comes in an adversarial or in a random order. They designed an optimal algorithm for the adversarial arrivals and a near optimal algorithm for the random arrivals. Duetting et al. [13] generalized the latter setting to secretaries with advice and found an optimal algorithm for the random arrival variant. Correa et al. [6] proposed a slightly different model in which each value is sampled independently with probability p and designed optimal algorithms for all p. This model bridges the secretary problem (when $p = 0$) and the i.i.d. prophet inequality (when $p = 1$). These models have similar flavour to our problem, but are not directly comparable.

Another line of research in sample-based prophet inequality studies the sample complexity, i.e., how many samples are needed to almost (up to an ε error) match the competitive ratio in the full-information case. First, Correa et al. [7] proved that $O(n^2)$ samples are sufficient to get the competitive ratio of $0.745-\varepsilon$ in the i.i.d. prophet inequality setting, where the optimal algorithm [9] with full information is 0.745-competitive. The sample complexity was later improved to $O(n/\varepsilon^6)$ by Rubinstein, Wang, and Weinberg [26]. Guo et al. [19] further improved the dependency on ε by establishing an upper bound of $O(n/\varepsilon^2)$.

2 Preliminaries

p-Sample Prophet Inequality There are n boxes, whose values $\mathbf{v} = (X_1, \ldots, X_n)$ are drawn independently from $\mathcal{D}_1 \times \ldots \times \mathcal{D}_n$. In contrast to the classic prophet inequality, the algorithm does not have knowledge about the underlying distribution in advance. Instead, for each random variable X_i we observe a sample $\hat{X}_i$ independently for all $i \in [n]$ with probability p. For simplicity of notations, we assume $\hat{X}_i = 0$ when we do not see a sample. The goal is to maximize the expected value of the accepted box and to compete against the expected maximum $\mathbb{E}[\max_i X_i]$.

Our algorithm is defined as the following.

Definition 2.1 (Max-Sample algorithm)**.** *Given as input samples $\hat{X}_1, \ldots, \hat{X}_n$, let $T = \max_{i\in[n]} \hat{X}_i$ and $i^* \overset{\text{def}}{=\!=} \operatorname{argmax}_i \hat{X}_i$ ($i^* = 0$ if $T = 0$). Let X_j be the first observed random variable exceeding T.*

$$\textit{If } j \neq i^*, \ \textit{take } X_j, \qquad \textit{If } j = i^* \ \textit{then} \begin{cases} \textit{take } X_j & \textit{w.p. } \frac{2p}{1+p} \\ \textit{skip } X_j, \ \textit{take next } X_\ell > T & \textit{w.p. } \frac{1-p}{1+p} \end{cases}$$

Theorem 2.1. *Max-Sample is $\frac{p}{1+p}$-competitive for the p-sample prophet inequality problem. Moreover, the ratio is the best possible for any $p \in [0, 1]$.*

For the algorithmic part of the result, we shall focus on the following card model and analyze the performance of our algorithm. The card model is adapted from the work of Rubinstein, Wang, and Weinberg [26] and Correa et al. [5].

Card Model. Each box corresponds to a card $C_i = \{a_i, b_i\}$ with two unknown values written on either side. Each card is put on the table with one of its sides independently and uniformly at random facing down and the other side facing up. The card i corresponds to an ordered pair (v_i, s_i): a value at the bottom, and a value on top. I.e., $\mathbf{Pr}[v_i = a_i, s_i = b_i] = \mathbf{Pr}[v_i = b_i, s_i = a_i] = 0.5$. The online algorithm gets to see some of the top values in the initial stage before making any decisions. Each top value s_i is revealed (independently for all $i \in [n]$) to the algorithm with probability p, with remaining probability $1-p$ the value is erased and is substituted with a blank. The online algorithm proceeds by flipping the cards one by one starting from C_1 and until the last card C_n. After flipping a card C_i, the algorithm observes the value v_i at the bottom and may either take it and stop, or discard the card C_i and continue. We denote the set of cards with revealed samples as $R \subseteq [n]$ and the distribution of the revealed samples $R \sim \mathcal{R}$. Also for each card $i \in [n]$ we use $r_i \in \{0,1\}$ to indicate whether the sample on top of card i is revealed ($r_i = 1$) or not ($r_i = 0$). We denote by $\mathbf{r} \in \{0,1\}^n$ the vector of revealed samples. We slightly abuse the notations and use $\mathcal{R}$ to denote the distribution of $\mathbf{r} \sim \mathcal{R}$. The algorithm sees revealed samples $\mathbf{s}(R)$ and aims to maximize the value of the accepted card and compete against the prophet in the card model: $\mathop{\mathbb{E}}_{\mathbf{v},\mathbf{s}}[\max_{i\in[n]} v_i]$.

For analysis purpose, we sort the multi-set of values $V = \{a_i\}_{i=1}^n \cup \{b_i\}_{i=1}^n$ in decreasing order. We denote the elements in the sorted multi-set V as $w_1 \geq w_2 \geq \ldots \geq w_{2n}$. We use $\sigma : [2n] \to [n]$ to denote the indexes of the cards in V. Specifically, σ_1 is the index of the card with the largest values in V, σ_2 is the index of the card with the second largest value, etc.

It is straightforward to observe that a competitive algorithm for the card model preserves its competitive ratio in the p-sample prophet inequality setting, by setting the values $\{a_i, b_i\}$ to be independent samples of $\mathcal{D}_i$.

Roadmap. In Section 3, we consider the simpler task of stopping at the maximum value card. Built on it, we prove our main result in Section 4. Finally, in Section 5, we provide a matching hardness result.

3 Stopping at the maximum

Consider the case when the largest value w_1 is much larger than the rest $w_i \in V$. In this case, the contribution of all other values to the expected reward of our algorithm and the prophet are negligibly small and the question is how often our algorithm stops at C_{σ_1} and gets $v_{\sigma_1} = w_1$.

Our objective then is to stop at the global maximum w_1 in V. The prophet gets w_1 with probability 0.5, whenever $v_{\sigma_1} = w_1$, i.e., when w_1 is at the bottom of the card C_{σ_1}. We show that Max-Sample stops at the maximum w_1 with probability at least $\frac{p}{2(1+p)}$, which gives us the desired guarantee in the special case when w_1 is much larger than all other $w_i \in V$. The analysis will be helpful for obtaining the result in general case.

Theorem 3.1. *Given that maximum w_1 is on the value side, i.e., $v_{\sigma_1} = w_1$, Max-Sample stops at the maximum with probability at least $\frac{p}{1+p}$.*

Before we proceed with the proof, we give an example demonstrating why the original algorithm of Rubinstein et al. [26] of accepting the first item above maximum sample has strictly worse performance than $\frac{p}{1+p}$. The instance has $n = 2$ boxes: the first box with distribution $F_1 = \texttt{Uni}[1,2]$, the second box with distribution $F_2 = \{v = 10000 \text{ w.p. } \frac{1}{100}, 0 \text{ w.p. } \frac{99}{100}\}$; let $p = \frac{1}{2}$. We may only consider the case when $X_2 = 10000, \hat{X}_2 = 0$ that contributes $10000 \cdot \frac{1}{100} \cdot \frac{199}{200} = 99.5$ to the expected value of the prophet, since the total contribution in all other cases is less than 3. In this case the algorithm of Rubinstein et al. gets X_1, if and only if $\hat{X}_1 > X_1$ which happens with probability $\frac{p}{2} < \frac{p}{1+p}$.

Proof. One difficulty in the analysis is that we know neither the order of cards, nor the pairings of $w_1, w_2, \ldots, w_{2n}$ (i.e., which pairs of them are on the same cards). Our approach in dealing with so many possibilities will be to describe a sequence of disjoint events that guarantee our algorithm to stop at w_1. We still need to consider a few cases, but only a small number.

We begin constructing these events by considering the cards with the largest values $w_2, \dots, w_t$ until the next one $\sigma_{t+1} \in \{\sigma_1, \dots, \sigma_t\}$, i.e., the first time when w_{t+1} is on the same card with one of the previous values $\{w_1, \dots, w_t\}$. Let us first deal with the case when $\sigma_{t+1} \neq \sigma_1$, i.e., w_{t+1} is on the same card with one of the $\{w_2, \dots, w_t\}$.

Case 1: $\sigma_{t+1} \neq \sigma_1$. We first consider the event $\mathcal{E}_1$ that w_2 is a visible sample ($s_{\sigma_2} = w_2$, $r_{\sigma_2} = 1$), then Max-Sample sets the threshold $T = w_2$ and waits until w_1 (recall that w_1 must be at the bottom of its card C_{σ_1}) at which point the algorithm must stop and take w_1. We have $\mathbf{Pr}[\mathcal{E}_1] = \frac{p}{2}$.

Next, if $s_{\sigma_2} = w_2$ and the sample w_2 is *not revealed* $r_{\sigma_2} = 0$, then we can look at w_3. If w_3 is a revealed sample ($w_3 = s_{\sigma_3}$, $r_{\sigma_3} = 1$) then Max-Sample must stop at w_1. This is our second event $\mathcal{E}_2$: ($s_{\sigma_2} = w_2, r_{\sigma_2} = 0$), and ($w_3 = s_{\sigma_3}, r_{\sigma_3} = 1$). Similarly, Max-Sample must stop at w_1 for each of the following events $\{\mathcal{E}_\ell\}_{\ell=1}^{t-1}$:

$$\mathcal{E}_\ell \stackrel{\text{def}}{=} \{\forall i \in [2, \ell] \quad (s_{\sigma_i} = w_i, r_{\sigma_i} = 0), \text{ and } (w_{\ell+1} = s_{\sigma_{\ell+1}}, r_{\sigma_{\ell+1}} = 1\})$$
$$\mathbf{Pr}\left[\mathcal{E}_\ell\right] = \left(\frac{1-p}{2}\right)^{\ell-1} \frac{p}{2}. \tag{1}$$

When we continue our sequence of events $\{\mathcal{E}_\ell\}_{\ell=1}^{t-1}$ to $\ell = t$, the value w_{t+1} appears on one of the previously fixed cards C_{σ_j} for $2 \leq j \leq t$. We note that Max-Sample algorithm skips the card with the maximum sample with probability $\frac{1-p}{1+p}$. Thus it may still stop at w_1 even when $v_{\sigma_j} = w_j$ (w_j is at the bottom of C_{σ_j} card). Finally, we define the last event $\mathcal{E}_t$ as follows:

$$\mathcal{E}_t \stackrel{\text{def}}{=} \left\{\forall i \in [t] \setminus \{1, j\} \quad \begin{pmatrix} s_{\sigma_i} = w_i \\ r_{\sigma_i} = 0 \end{pmatrix} \text{ and } \begin{pmatrix} w_j = v_{\sigma_j}, & w_{t+1} = s_{\sigma_j} \\ r_{\sigma_j} = 1, & \text{alg. ignores } w_j \end{pmatrix}\right\}$$
$$\mathbf{Pr}\left[\mathcal{E}_t\right] = \left(\frac{1-p}{2}\right)^{t-2} \cdot \frac{p}{2} \cdot \frac{1-p}{1+p}. \tag{2}$$

As all events $\{\mathcal{E}_\ell\}_{\ell=1}^{t}$ are disjoint, we may combine (1) and (2) and get

$$\mathbf{Pr}\left[\text{alg. takes } w_1\right] \geq \sum_{\ell=1}^{t} \mathbf{Pr}\left[\mathcal{E}_\ell\right] = \frac{p}{2} \cdot \left[\left(\frac{1-p}{2}\right)^{t-2} \cdot \frac{1-p}{1+p} + \sum_{i=0}^{t-2} \left(\frac{1-p}{2}\right)^{i}\right]$$
$$= \frac{p}{1+p}\left(\frac{1-p}{2}\right)^{t-1} + \frac{p}{2} \cdot \frac{1 - \left(\frac{1-p}{2}\right)^{t-1}}{1 - \left(\frac{1-p}{2}\right)} = \frac{p}{1+p} \tag{3}$$

Case 2: $\sigma_{t+1} = \sigma_1$. We have the same events $\{\mathcal{E}_\ell\}_{\ell=1}^{t-1}$ as in (1). The $\mathcal{E}_t$ is now a little different, as we want Max-Sample algorithm to stop at card $C_{\sigma_{t+1}}$:

$$\mathcal{E}_t \stackrel{\text{def}}{=} \left\{\forall i \in [t] \setminus \{1\} \quad \begin{pmatrix} s_{\sigma_i} = w_i \\ r_{\sigma_i} = 0 \end{pmatrix} \text{ and } \begin{pmatrix} w_1 = v_{\sigma_1}, & w_{t+1} = s_{\sigma_1} \\ r_{\sigma_j} = 1, & \text{alg. takes } w_1 \end{pmatrix}\right\}$$
$$\mathbf{Pr}\left[\mathcal{E}_t\right] = \left(\frac{1-p}{2}\right)^{t-2} \cdot p \cdot \frac{2p}{1+p}. \tag{4}$$

We continue the sequence of events $\{\mathcal{E}_\ell\}_{\ell=1}^{t}$ after t by considering new cards $C_{\sigma_{t+2}}, C_{\sigma_{t+3}}, \dots, C_{\sigma_k}$ until we get $\sigma_{k+1} \in \{\sigma_1, \dots, \sigma_k\} \setminus \{\sigma_1, \sigma_{t+1}\}$, i.e., the first time w_{k+1} appears on the same card with one of the previous $\{w_1, \dots, w_k\}$. Notice that w_1 and w_{t+1} are on the same card, and w_1 is always at the bottom of C_{σ_1}. We would like the sample $w_{t+1} = s_{\sigma_1}$ not to be revealed (i.e., $r_{\sigma_1} = 0$), which happens with probability $(1-p)$. We define $\{\mathcal{E}_\ell\}_{\ell=t+1}^{k-1}$ as follows

$$\mathcal{E}_\ell \stackrel{\text{def}}{=\!=} \left\{ \begin{pmatrix} v_{\sigma_1} = w_1 \\ r_{\sigma_1} = 0 \end{pmatrix}, \forall i \in [\ell] \setminus \{1, t+1\} \; \begin{pmatrix} s_{\sigma_i} = w_i \\ r_{\sigma_i} = 0 \end{pmatrix}, \text{ and } \begin{pmatrix} w_\ell = s_{\sigma_\ell} \\ r_{\sigma_\ell} = 1 \end{pmatrix} \right\}$$

$$\mathbf{Pr}\left[\mathcal{E}_\ell\right] = (1-p) \cdot \left(\frac{1-p}{2}\right)^{\ell-2} \cdot \frac{p}{2}. \tag{5}$$

Finally, let j be the index such that w_{k+1} appears on one of the previously fixed cards C_{σ_j} for $2 \le j \le k$. We define the last event $\mathcal{E}_k$ similar to (2) as follows.

$$\mathcal{E}_k \stackrel{\text{def}}{=\!=} \left\{ \forall i \in [k] \setminus \{j, t+1\} \; \begin{pmatrix} s_{\sigma_i} = w_i \\ r_{\sigma_i} = 0 \end{pmatrix}, \quad \begin{pmatrix} w_j = v_{\sigma_j}, w_{k+1} = s_{\sigma_j} \\ r_{\sigma_j} = 1, \text{ alg. ignores } w_j \end{pmatrix} \right\}$$

$$\mathbf{Pr}\left[\mathcal{E}_k\right] = (1-p) \cdot \left(\frac{1-p}{2}\right)^{k-3} \cdot \frac{p}{2} \cdot \frac{1-p}{1+p}. \tag{6}$$

As all events $\{\mathcal{E}_\ell\}_{\ell=1}^{k}$ are disjoint, we combine (1),(4), (5), and (6) to get

$$\begin{aligned}
\mathbf{Pr}\left[\text{alg. takes } w_1\right] &\ge \sum_{\ell=1}^{k} \mathbf{Pr}\left[\mathcal{E}_\ell\right] = \frac{p}{2} \cdot \sum_{i=0}^{t-2} \left(\frac{1-p}{2}\right)^i + \left(\frac{1-p}{2}\right)^{t-2} \cdot \frac{2p^2}{1+p} + \\
& p \cdot \sum_{\ell=t+1}^{k-1} \left(\frac{1-p}{2}\right)^{\ell-1} + \left(\frac{1-p}{2}\right)^{k-1} \cdot \frac{2p}{1+p} > \frac{p}{1+p}\left[1 - \left(\frac{1-p}{2}\right)^{t-1}\right] \\
& + \frac{p^2}{1+p}\left(\frac{1-p}{2}\right)^{t-1} + p \cdot \left(\frac{1-p}{2}\right)^{t} \frac{1 - \left(\frac{1-p}{2}\right)^{k-1-t}}{1 - \frac{1-p}{2}} + \left(\frac{1-p}{2}\right)^{k-1} \cdot \frac{2p}{1+p} \\
& = \frac{p}{1+p} - \frac{p}{1+p}\left(\frac{1-p}{2}\right)^{t-1}(1-p) + \frac{2p}{1+p}\left(\frac{1-p}{2}\right)^{t} \cdot \left[1 - \left(\frac{1-p}{2}\right)^{k-t-1}\right] + \\
& \left(\frac{1-p}{2}\right)^{k-1} \frac{2p}{1+p} = \frac{p}{1+p},
\end{aligned} \tag{7}$$

where to get the last inequality we simply decreased the term $\mathbf{Pr}[\mathcal{E}_t]$ in (4) to $\left(\frac{1-p}{2}\right)^{k-1} \frac{p^2}{1+p}$.

Theorem 3.1 follows from (3) in the case $\sigma_{t+1} \neq \sigma_1$ and from (7) in the case $\sigma_{t+1} = \sigma_1$. □

4 Maximizing Expectation

In this section we prove our main result that Max-Sample algorithm achieves optimal competitive ratio of $\frac{p}{1+p}$ on arbitrary instances. To this end we consider a few special instances with top k values being almost equal to each other[1] and much larger than the remaining w_i for $i \in [2n] \setminus [k]$. We call it a top-k instance for any $k \le 2n$. It turns out that restricting our attention only to the top-k instances is without loss of generality for the Max-Sample algorithm. We prove next that Max-Sample is a $\frac{p}{1+p}$ approximation to the prophet on any top-k instance for each $k \in [2n]$ using similar approach to what we did in Section 3 for the top-1 instances, but with a more elaborate case analysis.

[1] E.g., $w_1 = w_2 + \varepsilon = \ldots = w_k + (k-1)\varepsilon$, for some negligibly small $\varepsilon > 0$.

Theorem 4.1. *Max-Sample algorithm is a $\frac{p}{1+p}$-approximation to the prophet for any value of $p \in (0, 1]$.*

Proof. We analyse Max-Sample in the card model and first show that restricting our attention only to the top-k instances is without loss of generality.

Claim 4.1. *Suppose a* rank-based *ALG (i.e., ALG only uses comparisons "$>, <$" between variables and samples) is an $\alpha < 1$ approximation to the prophet on any top-k instance in the card model for each $k \geq 1$. Then ALG is an α-approximation to the prophet on every instance.*

Proof. The fact that ALG is an α-approximation to the prophet on a top-k instance means that

$$\mathbf{Pr}_{\mathbf{v},\mathbf{r}}\left[\mathsf{ALG}(\mathbf{v},\mathbf{r}) \text{ gets } w_i \text{ for an } i \in [k]\right] \geq \alpha \cdot \mathbf{Pr}_{\mathbf{v}}\left[\exists i \in [k] \; v_{\sigma_i} = w_i\right] \tag{8}$$

for this instance. As ALG is an ordinal algorithm the same guarantee holds for *any instance* that is not necessarily a top-k. Expected performance of the ALG can be written as

$$\begin{aligned}\mathbb{E}\left[\mathsf{ALG}\right] = \sum_{k=1}^{2n} w_k \cdot \mathbf{Pr}\left[\mathsf{ALG} \text{ gets } w_k\right] = \sum_{k=1}^{2n} w_k \cdot \Big(\mathbf{Pr}\left[\mathsf{ALG} \text{ gets } w_i, i \in [k]\right] \\ - \mathbf{Pr}\left[\mathsf{ALG} \text{ gets } w_i, i \in [k-1]\right]\Big) = \sum_{k=1}^{2n} \mathbf{Pr}\left[\mathsf{ALG} \text{ gets } w_i, i \in [k]\right] \cdot (w_k - w_{k+1}) \\ \geq \alpha \cdot \sum_{k=1}^{2n} \mathbf{Pr}\left[\exists i \in [k] \; v_{\sigma_i} = w_i\right] \cdot (w_k - w_{k+1}) = \alpha \cdot \mathbb{E}\left[\mathsf{Prophet}\right],\end{aligned}$$

where $\mathbf{Pr}[\mathsf{ALG} \text{ gets } w_i, i \in [0]] = w_{2n+1} = 0$; we used (8) to get the inequality. □

To conclude the proof of Theorem 4.1, we only need to show that Max-Sample is a $\frac{p}{1+p}$-approximation to the prophet on a top-k instance for each $k \in [2n]$. Section 3 already gives the desired result for $k = 1$. For $k \geq 2$ we consider the sequence of cards $(\sigma_i)_{i=1}^k$ with the top k values. There are two cases: 1) there is a pair of w_i, w_j on the same card ($\sigma_i = \sigma_j$), or 2) all $(\sigma_i)_{i=1}^k$ are different.

Case 1. $\exists \; \sigma_j = \sigma_i, \; i, j \in [k]$. Let us consider the first time two top values appear on the same card, i.e., the smallest $i \leq k$ with $\sigma_i = \sigma_j$ for a $j < i$. Notice that the prophet can get one of the top-i values (either w_i or w_j), i.e., $\mathbf{Pr}[\exists j \in [i] \; v_{\sigma_j} = w_j] = 1$. On the other hand, for the Max-Sample it is only harder to stop at one of the top-i values. Hence, we can assume without loss of generality that $k = i$ and that $\sigma_j = \sigma_k$ is the only two top-k values on the same card. We distinguish three cases based on the index j.

Case 1.a. $k = 2$. Then $\sigma_1 = \sigma_2$, i.e., w_1 and w_2 are on the same card. Then consider the event $\mathcal{E}_0$ that $(v_{\sigma_1} = w_1, s_{\sigma_1} = w_2, r_{\sigma_1} = 1)$, then Max-Sample succeeds with probability $\frac{2p}{1+p}$. On the other hand, if $r_{\sigma_1} = 0$ (sample $s_{\sigma_1} \in \{w_1, w_2\}$ is not revealed) we can use the same events $(\mathcal{E}_i)_{i \geq 1}$ from Section 3 (count starts from w_3 instead of w_2) to guarantee that Max-Sample stops at w_1 or w_2 (whichever is at the bottom of C_{σ_1}). Overall, the events $(\mathcal{E}_i)_{i \geq 0}$ give us the desired guarantee

$$\mathbf{Pr}\left[\text{alg. wins}\right] \geq \frac{p}{2} \cdot \frac{2p}{1+p} + (1-p) \cdot \sum_{i \geq 1} \mathbf{Pr}\left[\mathcal{E}_i\right] \geq \frac{p^2}{1+p} + (1-p)\frac{p}{1+p} = \frac{p}{1+p}.$$

Case 1.b. $k > 2$ and $\sigma_j = \sigma_k \neq \sigma_1$. The Max-Sample wins in the event $\mathcal{E}_0 : (v_{\sigma_1} = w_1, r_{\sigma_k} = 1)$. On the other hand, when $(v_{\sigma_1} = w_1, r_{\sigma_k} = 0)$, we can use the same events $(\mathcal{E}_i)_{i\geq 1}$ as in Section 3 with a small modification that w_j, w_k and their respective card C_{σ_j} are removed from the sequence $(w_i)_{i=1}^{2n}$ to guarantee the win of Max-Sample. Indeed, the card C_{σ_j} may only cause the algorithm to stop early at w_j or w_k, which is a win for the algorithm. We have

$$\mathbf{Pr}\left[\text{alg. wins}\right] \geq \mathbf{Pr}\left[\mathcal{E}_0\right] + \frac{1}{2}(1-p) \cdot \sum_{i\geq 1} \mathbf{Pr}\left[\mathcal{E}_i\right] \geq \frac{p}{2} + \frac{1-p}{2}\frac{p}{1+p} = \frac{p}{1+p}.$$

Case 1.c. $k > 2$ and $\sigma_k = \sigma_1$. We consider first what happens with the card C_{σ_1}. First, let us consider what happens when $(s_{\sigma_1} = w_k, r_{\sigma_1} = 0)$. The Max-Sample wins if at least one of the top values $w_i, i \in [2, k]$ is revealed as a sample $(s_{\sigma_i} = w_i, r_{\sigma_i} = 1)$. We define this event $\mathcal{E}_I$ as

$$\mathcal{E}_I \stackrel{\text{def}}{=\!=} \{(s_{\sigma_1} = w_k, r_{\sigma_1} = 0),\ \exists\ 1 < i < k\ (s_{\sigma_i} = w_i, r_{\sigma_i} = 1)\}$$
$$\mathbf{Pr}\left[\mathcal{E}_I\right] = \frac{1-p}{2} \cdot \left(1 - \left(1 - \frac{p}{2}\right)^{k-2}\right) \tag{9}$$

Next, let us consider what happens when $(v_{\sigma_1} = w_1, s_{\sigma_1} = w_k, r_{\sigma_1} = 1)$. The algorithm is guaranteed to win when one of the w_i for $1 < i < k$ appears at the bottom, or as a revealed sample. The Max-Sample also wins when all of the w_i for $1 < i < k$ appear as a hidden samples and Max-Sample decides not to skip w_1 when it reaches $s_{\sigma_1} = w_k$. Formally, we define these two events $\mathcal{E}_{II}, \mathcal{E}_{III}$ as

$$\mathcal{E}_{II} \stackrel{\text{def}}{=\!=} \{(s_{\sigma_1} = w_k, r_{\sigma_1} = 1),\ \exists\ 1 < i < k\ (v_{\sigma_i} = w_i \text{ or } s_{\sigma_i} = w_i, r_{\sigma_i} = 1)\}$$
$$\mathcal{E}_{III} \stackrel{\text{def}}{=\!=} \{(s_{\sigma_1} = w_k, r_{\sigma_1} = 1),\ \forall\ 1 < i < k\ (s_{\sigma_i} = w_i, r_{\sigma_i} = 0), \mathsf{ALG} \text{ takes } w_1\}$$
$$\mathbf{Pr}\left[\mathcal{E}_{II} \sqcup \mathcal{E}_{III}\right] = \frac{p}{2} \cdot \left(1 - \left(\frac{1-p}{2}\right)^{k-2} + \left(\frac{1-p}{2}\right)^{k-2} \cdot \frac{2p}{1+p}\right) \tag{10}$$

Finally, let us consider what happens when w_k is at the bottom of the card C_{σ_1} and w_1 is not revealed as a sample $(v_{\sigma_1} = w_k, r_{\sigma_1} = 0)$. We would like to treat w_k as w_1 from Section 3 and construct events that guarantee Max-Sample to stop at w_k. The main problem is that if any of the w_i for $i \in [2, k-1]$ appears as a revealed sample, then the algorithm will never stop at w_k. To avoid this issue we will add the condition that no w_i for $i \in [2, k-1]$ is revealed as a sample (note that if w_i appears at the bottom of C_{σ_i}, it can only help Max-Sample to win). We use the events $\{\mathcal{E}_\ell\}_{\ell\geq 1}$ from Section 3 with a modification that w_k plays the role of w_1 and all $w_2, \ldots, w_{k-1}$ are ignored or equivalently treated as very small numbers (i.e., $(w_i)_{i\geq 2}$ in Section 3 correspond to $(w_i)_{i\geq k+1}$ in our instance, and w_1 in Section 3 corresponds to w_k here)[2]. Notice that if a card with $w_i, i \in [2, k-1]$ is used in an event $\mathcal{E}_\ell$, then the other value $w_j, j \geq k+1$ on the card C_{σ_i} must be a sample $(s_{\sigma_j} = w_j, v_{\sigma_j} = w_i)$. I.e., we do not need to worry that w_i is revealed as a sample. Thus for each event $\mathcal{E}_\ell$ the algorithm wins in the event

$$\mathcal{E}'_\ell \stackrel{\text{def}}{=\!=} \{(v_{\sigma_1} = w_k, r_{\sigma_1} = 0) \wedge \mathcal{E}_\ell \wedge \{\forall 1 < i < k\ v_{\sigma_i} = w_i \vee (s_{\sigma_i} = w_i, r_{\sigma_i} = 0)\}\}$$
$$\mathbf{Pr}\left[\mathcal{E}'_\ell\right] \geq \frac{1-p}{2} \cdot \left(1 - \frac{p}{2}\right)^{k-2} \cdot \mathbf{Pr}\left[\mathcal{E}_\ell\right]. \tag{11}$$

[2] We can assume that once w_i is set to 0 for $1 < i < k$, it is small enough to not appear in the $\mathcal{E}_\ell$. Indeed, we can add a few dummy cards with both sides having negligibly small numbers in the beginning of the sequence that do not affect performance of Max-Sample, but still bigger than $w_i \leftarrow 0$.

When we combine the events defined by (9), (10),(11) we get

$$\mathbf{Pr}\left[\text{alg. wins}\right] \geq \mathbf{Pr}\left[\mathcal{E}_I \sqcup \mathcal{E}_{II} \sqcup \mathcal{E}_{III} \sqcup \bigsqcup_{\ell\geq 1} \mathcal{E}'_\ell\right] \geq \frac{1-p}{2}\cdot\left(1-\left(1-\frac{p}{2}\right)^{k-2}\right)+$$

$$\frac{p}{2}\cdot\left(1-\left(\frac{1-p}{2}\right)^{k-2}+\left(\frac{1-p}{2}\right)^{k-2}\cdot\frac{2p}{1+p}\right)+\frac{1-p}{2}\left(1-\frac{p}{2}\right)^{k-2}\cdot\frac{p}{1+p}$$

$$=\frac{1-p}{2}\cdot\left(1-\frac{1}{1+p}\left(1-\frac{p}{2}\right)^{k-2}\right)+\frac{p}{2}\cdot\left(1-\frac{1-p}{1+p}\left(\frac{1-p}{2}\right)^{k-2}\right)\overset{k=3}{>}$$

$$\frac{1-p}{2}\cdot\left(1-\frac{1}{1+p}\left(1-\frac{p}{2}\right)\right)+\frac{p}{2}\cdot\left(1-\frac{1-p}{1+p}\right)=\frac{p}{1+p},$$

where to get the last inequality we used that the previous expression is minimized for $k=3$ (recall that $k>2$) and also that $\left(\frac{1-p}{2}\right)^{k-2}<1$. This concludes the proof for the case 1 as we have $\mathbf{Pr}[\text{alg. wins}]\geq\frac{p}{1+p}$ in each of the sub-cases 1.a, 1.b, and 1.c.

Case 2 $k\geq 2$ and $\forall\ 1\leq i<j\leq k\ \sigma_i\neq\sigma_j$. The main difficulty in this case is that the value of the prophet $\mathbf{Pr}[\exists\ j\in[i]\ v_{\sigma_j}=w_j]=1-\frac{1}{2^k}$, which depends on k. On the positive side, there are no sub-cases here unlike case 1. We consider first what happens when at least one of the $w_i, i\in[k]$ is revealed as a sample. Let $j\in[k]$ be the first w_j with $(s_{\sigma_j}=w_j, r_{\sigma_j}=1)$. For each $j\geq 2$, Max-Sample algorithm wins when at least one of the $w_i, i<j$ appears at the bottom of its card C_{σ_i}. Formally, we define these events $(\mathcal{E}^*_j)^k_{j\geq 2}$ as

$$\mathcal{E}^*_j \overset{\text{def}}{=\!=} \left\{(s_{\sigma_j}=w_j, r_{\sigma_j}=1),\quad \begin{array}{ll}\forall\ i<j & (v_{\sigma_i}=w_i\ \vee\ r_{\sigma_i}=0)\\ \text{not}\ \forall\ i<j & (s_{\sigma_i}=w_i\ \wedge\ r_{\sigma_i}=0)\end{array}\right\}$$

$$\mathbf{Pr}\left[\mathcal{E}^*_j\right]=\frac{p}{2}\cdot\left(\left(1-\frac{p}{2}\right)^{j-1}-\left(\frac{1-p}{2}\right)^{j-1}\right) \tag{12}$$

Next, consider the event $\mathcal{E}'$ that none of $w_i, i\in[k]$ is revealed as a sample and at least one of them is at the bottom of its card C_{σ_i}. In this case, we need to consider other $(w_i)_{i\geq k+1}$ to guarantee the win of Max-Sample. To this end, we would like to use the events $(\mathcal{E}_\ell)_{\ell\geq 1}$ from Section 3 with the following modification: all $w_1,\ldots,w_k$ are ignored, i.e., $(w_i)_{i\geq 2}$ from Section 3 correspond to $(w_i)_{i\geq k+1}$ in our instance. If event $\mathcal{E}_\ell$ does not specify position of any of the cards $C_{\sigma_1},\ldots,C_{\sigma_k}$, then Max-Sample wins in the event $\mathcal{E}'_\ell$ defined as:

$$\mathcal{E}'_\ell \overset{\text{def}}{=\!=} \left\{\mathcal{E}_\ell\ ,\quad \begin{array}{ll}\forall\ i\in[k] & (v_{\sigma_i}=w_i\ \vee\ r_{\sigma_i}=0)\\ \text{not}\ \forall\ i\in[k] & (s_{\sigma_i}=w_i\ \wedge\ r_{\sigma_i}=0)\end{array}\right\}$$

$$\mathbf{Pr}\left[\mathcal{E}'_\ell\right]=\left(\left(1-\frac{p}{2}\right)^{k}-\left(\frac{1-p}{2}\right)^{k}\right)\cdot\mathbf{Pr}\left[\mathcal{E}_\ell\right] \tag{13}$$

Now, if $\mathcal{E}_\ell$ specifies the position of any of the cards $C_{\sigma_1},\ldots,C_{\sigma_k}$, then let us consider the first time $j\geq k+1$ when $\sigma_j=\sigma_i$, for an $i\in[k]$. Note that in this case w_i must be at the bottom of C_{σ_j} $(v_{\sigma_i}=w_i)$. We can treat w_i as w_1 in the event $\mathcal{E}_\ell$ from Section 3 and ignore the remaining $w_t, t\in[k]\setminus\{i\}$. Then for every card $C_{\sigma_t}, t\in[k]$ that is specified in $\mathcal{E}_\ell$, we have $v_{\sigma_t}=w_t$. We immediately get $\neg\forall\ i\in[k]\ (s_{\sigma_i}=w_i\wedge r_{\sigma_i}=0)$ and specifically for the card C_{σ_t} we get

$(v_{\sigma_t} = w_t \ \vee \ r_{\sigma_t} = 0)$. We still need to check that $v_{\sigma_t} = w_t \ \vee \ r_{\sigma_t} = 0$ for the cards not specified in $\mathcal{E}_\ell$. Thus, for the event $\mathcal{E}'_\ell$ formally defined in (13) we get

$$\mathbf{Pr}\left[\mathcal{E}'_\ell\right] \geq \left(1 - \frac{p}{2}\right)^{k-1} \cdot \mathbf{Pr}\left[\mathcal{E}_\ell\right] \geq \left(\left(1 - \frac{p}{2}\right)^k - \left(\frac{1-p}{2}\right)^k\right) \cdot \mathbf{Pr}\left[\mathcal{E}_\ell\right] \tag{14}$$

Finally, we combine the events $\{\mathcal{E}^*_j\}_{j\geq 2}^k$ and $\{\mathcal{E}'_\ell\}_{\ell\geq 1}$ and use (12),(14) to get

$$\mathbf{Pr}\left[\text{alg. wins}\right] \geq \mathbf{Pr}\left[\bigsqcup_{j=2}^{k} \mathcal{E}^*_j \sqcup \bigsqcup_{\ell\geq 1} \mathcal{E}'_\ell\right] \geq \frac{p}{2} \cdot \sum_{j=1}^{k-1} \left(\left(1 - \frac{p}{2}\right)^j - \left(\frac{1-p}{2}\right)^j\right) +$$
$$\left(\left(1 - \frac{p}{2}\right)^k - \left(\frac{1-p}{2}\right)^k\right) \cdot \sum_{\ell\geq 1} \mathbf{Pr}\left[\mathcal{E}_\ell\right] \geq \frac{p}{2} \cdot \sum_{j=1}^{k-1} \left(\left(1 - \frac{p}{2}\right)^j - \left(\frac{1-p}{2}\right)^j\right) +$$
$$\left(\left(1 - \frac{p}{2}\right)^k - \left(\frac{1-p}{2}\right)^k\right) \cdot \frac{p}{1+p} = \frac{p}{2}\left(1 - \frac{p}{2}\right) \frac{1 - \left(1 - \frac{p}{2}\right)^{k-1}}{p/2}$$
$$- \frac{p}{2}\left(\frac{1-p}{2}\right) \frac{1 - \left(\frac{1-p}{2}\right)^{k-1}}{1 - \frac{1-p}{2}} + \left(\left(1 - \frac{p}{2}\right)^k - \left(\frac{1-p}{2}\right)^k\right) \frac{p}{1+p} =$$
$$1 - \frac{p}{2} - \frac{1}{1+p}\left(1 - \frac{p}{2}\right)^k - \frac{p(1-p)}{2(1+p)} = \frac{1}{1+p}\left(1 - \left(1 - \frac{p}{2}\right)^k\right) \tag{15}$$

We are left to verify that the right hand side of (15) is at least $\frac{p}{1+p} \cdot \mathsf{Prophet} = \frac{p}{1+p} \cdot \left(1 - \frac{1}{2^k}\right)$. This is equivalent to showing that $f(p) \stackrel{\text{def}}{=\!=} 1 - \left(1 - \frac{p}{2}\right)^k \geq g(p) \stackrel{\text{def}}{=\!=} p \cdot (1 - \frac{1}{2^k})$. Now, observe that $f(0) = g(0)$, $f(1) = g(1)$, and $f'(p) - g'(p) = \frac{k}{2}\left(1 - \frac{p}{2}\right)^{k-1} - 1 + \frac{1}{2^k}$ is a decreasing function in p that is positive at $p = 0$ (recall that $k \geq 2$). These three conditions imply that $f(p) - g(p) \geq 0$ for any $p \in [0, 1]$. □

5 Matching Lower Bound

We give in this section a matching lower bound of $\frac{p}{1+p}$. Interestingly, our construction has the property that the maximum value among all n values together with all n samples (revealed or not) is almost surely much larger than the rest $2n - 1$ numbers, that is the upper bound of $\frac{p}{1+p}$ from section 3 is tight.

Our construction is as follows for any fixed constant $p \in (0, 1]$.

Example 5.1. *Set $\varepsilon = o(p) > 0$. Let the number of variables $n = \Theta\left(\frac{1}{\varepsilon^2}\right)$. Define the distribution $F_0 = \{v = 0 \text{ w.p. } 1\}$ and distributions $F_i \stackrel{\text{def}}{=\!=} \{v = \frac{1}{\varepsilon^i} \text{ w.p. } \varepsilon, v = 0 \text{ w.p. } 1 - \varepsilon\}$ for all $i \in [n]$. We construct the following mixture of n instances $\{I_i\}_{i=1}^n$.*

$$i\text{-th instance } I_i : \forall j \leq i \ \ \mathcal{D}_j = F_j, \forall j > i \ \ \mathcal{D}_j = F_0 \quad \mathbf{Pr}\left[I_i\right] = \frac{\varepsilon^{i-1}}{\sum_{j=0}^{n-1} \varepsilon^j}$$

The next two claims describe an optimal online algorithm ALG for this instance. Let $i^* \stackrel{\text{def}}{=\!=} \text{argmax}_i \hat{X}_i$ ($i^* = 1$ if all $\hat{X}_i = 0$).

Claim 5.1. *There is an optimal ALG that does not take any X_i with $i < i^*$.*

Proof. If ALG stops at any $i < i^*$, then its reward X_i is equal to or smaller than $\frac{1}{\varepsilon^{i^*-1}}$. On the other hand, ALG could wait until i^* and get a reward of at least $\frac{1}{\varepsilon^{i^*}}$ with probability ε, since the distribution $\mathcal{D}_{i^*} = F_i$. This gives at least as large expected reward of $\frac{1}{\varepsilon^{i^*}} \cdot \varepsilon$ as taking X_i. □

Claim 5.2. *The ALG that takes the first non zero X_i for $i \geq i^*$ is optimal.*

Proof. First, we may assume that ALG does not stop before i^* by Claim 5.1. Also we can assume that ALG skips any $X_i = 0$. Note that the revealed samples up until i^* give no information about the variables after i^*. Thus ALG should only consider zero samples after i^*, which we denote by a vector $\mathbf{s}_R$. Then for each $j \geq i^*$

$$\mathbf{Pr}_{I_i}\left[\mathbf{s}_R = \mathbf{0} \mid I_i = I_j\right] = (1-\varepsilon)^{f(j)}, \text{ where } f(j) \overset{\text{def}}{=\!=} |R \cap \{i^*, i^*+1, \ldots, j\}|.$$

Using Bayes rule and the law of total probability we can get

$$\mathbf{Pr}_{I_i}\left[I_i = I_j \mid i \geq i^*, \mathbf{s}_R = \mathbf{0}\right] = \frac{w_j}{\sum_{i \geq i^*} w_i}, \text{ where } w_j \overset{\text{def}}{=\!=} \varepsilon^{j-i^*} \cdot (1-\varepsilon)^{f(j)} \tag{16}$$

We will prove that an optimal ALG should always take $X_t = \frac{1}{\varepsilon^t}$ for any $t \geq i^*$ and any $\mathbf{s}_R$ by backward induction on t. The base of induction for $t = n$ is trivial. We prove induction step for a $t < n$ assuming that the induction hypothesis holds for all $t' : t < t' \leq n$. Assume towards a contradiction that an optimal algorithm ALG' does not take $X_t = \frac{1}{\varepsilon^t}$, then by the induction hypothesis ALG' must wait until the next variable $X_{t'} > 0$ and stop. Next, we will show that the expected reward in this case is strictly smaller than $\frac{1}{\varepsilon^t}$ – the reward ALG would have by stopping at X_t.

$$\begin{aligned}\mathbb{E}\left[\mathsf{ALG}'\right] &= \sum_{j>t}^{n} \frac{1}{\varepsilon^j} \cdot \mathbf{Pr}\left[\forall t < i < j \; X_i = 0, X_j > 0\right] \cdot \mathbf{Pr}_{I_i}\left[i \geq j \mid i \geq t, \mathbf{s}_R = \mathbf{0}\right] \\ &= \sum_{j>t}^{n} \frac{\varepsilon \cdot (1-\varepsilon)^{j-t-1}}{\varepsilon^j} \cdot \frac{\sum_{i \geq j}^{n} w_i}{\sum_{i \geq t}^{n} w_i} < \sum_{j>t}^{n} \frac{\varepsilon \cdot (1-\varepsilon)^{j-t-1}}{\varepsilon^j} \cdot \varepsilon^{j-t} < \frac{1}{\varepsilon^{t-1}} \sum_{i=0}^{\infty} (1-\varepsilon)^i \\ &= \frac{1}{\varepsilon^t} = \mathbb{E}\left[\mathsf{ALG}\right],\end{aligned} \tag{17}$$

where to get the first inequality we observe that $\varepsilon^{j-t} w_t \geq w_j$, $\varepsilon^{j-t} w_{t+1} \geq w_{j+1}, \ldots, \varepsilon^{j-t} w_{n-j+t} \geq w_n$ by formula (16) and thus $\varepsilon^{j-t} \cdot \sum_{i \geq t}^{n} w_i > \sum_{i \geq j}^{n} w_i$; in the second inequality we simply extended the range of summation from $i = n-t-1$ to infinity. The strict inequality (17) shows that ALG' cannot be optimal, and, therefore, an optimal ALG has to stop at X_t, which concludes the proof. □

Now, we can compare the optimal online algorithm described by Claim 5.2 with the prophet.

Theorem 5.1. *The competitive ratio of any online algorithm with respect to the prophet is at least $\frac{p}{1+p}$ for the Example 5.1.*

Proof. First, we get the following lower bound on the expected reward of the prophet

$$\mathsf{Prophet} \geq \sum_{\ell=1}^{n} \mathbf{Pr}_{I_i}\left[I_i = I_\ell\right] \cdot \mathbf{Pr}\left[X_\ell = \frac{1}{\varepsilon^\ell}\right] \cdot \frac{1}{\varepsilon^\ell} = \sum_{\ell=1}^{n} \frac{\varepsilon^{\ell-1} \cdot \varepsilon \cdot \frac{1}{\varepsilon^\ell}}{\sum_{i=0}^{n-1} \varepsilon^i} = n - o(n).$$

In what follows we get an upper bound on the expected reward ALG of the optimal online algorithm. Let us assume that the realized instance is $I_i = I_\ell$. We first observe that the total contribution from X_j with $1 \le j < \ell$ is not more than

$$\sum_{j=1}^{\ell-1} \mathbf{Pr}\left[X_j > 0\right] \cdot \frac{1}{\varepsilon^j} = \varepsilon \cdot \sum_{j=1}^{\ell-1} \varepsilon^{-j} = O\left(\frac{1}{\varepsilon^{\ell-2}}\right). \tag{18}$$

As we will see later this turns out to be a negligibly small amount. Next we get an upper bound on the probability that ALG stops at X_ℓ when $I_i = I_\ell$ and $X_\ell > 0$.

$$\mathbf{Pr}\left[\mathsf{ALG}\text{ takes } X_\ell \mid I_i = I_\ell, X_\ell > 0\right] = \mathbf{Pr}\left[\forall 1 \le i < \ell \; X_i = 0\right] + \sum_{j=1}^{\ell-1} \mathbf{Pr}\left[\forall j < i < \ell \; X_i = 0, X_j > 0\right] \cdot \mathbf{Pr}\left[\exists j < i \le \ell \; (\hat{X}_i > 0, r_i = 1)\right]$$

$$= (1-\varepsilon)^{\ell-1} + \sum_{i=0}^{\ell-2} \varepsilon \cdot (1-\varepsilon)^i \cdot (1-(1-p\varepsilon)^{i+1}) \tag{19}$$

We further estimate the term $A_\ell \stackrel{\text{def}}{=\!=} \sum_{i=0}^{\ell-2} \varepsilon \cdot (1-\varepsilon)^i \cdot (1-(1-p\varepsilon)^{i+1})$ in (19). We give an upper bound on A_ℓ by analysing a simple Markov chain M that corresponds to this summation. Markov chain M has 4 states $\{S, I, \mathsf{End}, \mathsf{Win}\}$; the random walk starts in the S state, and from there we can go either to Win with probability $p\varepsilon$, or to I with the remaining probability $1-p\varepsilon$; from state I we can either go back to S with probability $1-\varepsilon$, or go to End with remaining probability ε; finally, both states Win and End are terminal states, i.e., once the random walk gets in one of them, it stays there forever. The Win state represents that ALG successfully reaches X_ℓ and End state represents that ALG stops at an earlier random variable. Observe that

$$\mathbf{Pr}\left[\text{reach }\mathsf{Win}\right] = 1 - \mathbf{Pr}\left[\text{reach }\mathsf{End}\right] = 1 - \sum_{\ell=0}^{\infty} \varepsilon(1-\varepsilon)^\ell (1-p\varepsilon)^{\ell+1} = \sum_{\ell=0}^{\infty} \varepsilon(1-\varepsilon)^\ell - \sum_{\ell=0}^{\infty} \varepsilon(1-\varepsilon)^\ell (1-p\varepsilon)^{\ell+1} = \sum_{\ell=0}^{\infty} \varepsilon(1-\varepsilon)^\ell (1-(1-p\varepsilon)^{\ell+1}) \ge A_\ell$$

On the other hand, we have a simple recurrent equation for $\mathbf{Pr}[\text{reach }\mathsf{Win}] = p\varepsilon + (1-p\varepsilon)\cdot(1-\varepsilon)\cdot\mathbf{Pr}[\text{reach }\mathsf{Win}]$, which gives us

$$\frac{p+o(p)}{1+p} = \frac{p}{1+p-\varepsilon p} = \mathbf{Pr}\left[\text{reach }\mathsf{Win}\right] \ge A_\ell. \tag{20}$$

For the other term $B_\ell \stackrel{\text{def}}{=\!=} (1-\varepsilon)^{\ell-1}$ in (19) we will use that $n = \Omega(\frac{1}{\varepsilon^2})$ is rather large and thus for almost all ℓ the term B_ℓ is negligibly small. Now we can combine the bounds (18),(19), and (20) together to get the lower bound on expected reward ALG of the optimal online algorithm.

$$\mathsf{ALG} \leq \sum_{\ell=1}^{n} \mathbf{Pr}_{I_i}\left[I_i = I_\ell\right] \cdot \left[\frac{\mathbf{Pr}[X_\ell > 0]}{\varepsilon^\ell} \cdot \mathbf{Pr}\left[\text{ take } X_\ell \mid I_i = I_\ell, X_\ell > 0\right]\right.$$

$$\left. + O\left(\frac{1}{\varepsilon^{\ell-2}}\right)\right] = \sum_{\ell=1}^{n} \frac{\varepsilon^{\ell-1}}{\sum_{i=0}^{n-1}\varepsilon^i} \cdot \left[\frac{\varepsilon}{\varepsilon^\ell} \cdot \left(A_\ell + (1-\varepsilon)^{\ell-1}\right) + O\left(\frac{1}{\varepsilon^{\ell-2}}\right)\right] \leq$$

$$\sum_{\ell=1}^{n} \varepsilon^{\ell-1}\left[\frac{A_\ell}{\varepsilon^{\ell-1}} + \frac{(1-\varepsilon)^{\ell-1}}{\varepsilon^{\ell-1}} + O\left(\frac{1}{\varepsilon^{\ell-2}}\right)\right] < n \cdot A_n + O(n\varepsilon) + \sum_{\ell=1}^{\infty}(1-\varepsilon)^{\ell-1}$$

$$\leq \frac{np + o(np)}{1+p} + O(n\varepsilon) + \frac{1}{\varepsilon} = n \cdot \left(\frac{p}{1+p} + o(1)\right).$$

Combining this upper bound on ALG with a lower bound on the prophet we get the desired bound $\mathsf{ALG} \leq (1 + o(1))\frac{p}{1+p}\mathsf{Prophet}$. □

References

[1] Pablo Daniel Azar, Robert Kleinberg, and S. Matthew Weinberg. Prophet inequalities with limited information. In *SODA*, pages 1358–1377. SIAM, 2014.

[2] Moshe Babaioff, Yannai A. Gonczarowski, Yishay Mansour, and Shay Moran. Are two (samples) really better than one? In *EC*, page 175. ACM, 2018.

[3] Constantine Caramanis, Paul Dütting, Matthew Faw, Federico Fusco, Philip Lazos, Stefano Leonardi, Orestis Papadigenopoulos, Emmanouil Pountourakis, and Rebecca Reiffenhäuser. Single-sample prophet inequalities via greedy-ordered selection. In *SODA*, pages 1298–1325. SIAM, 2022.

[4] Shuchi Chawla, Jason D. Hartline, David L. Malec, and Balasubramanian Sivan. Multi-parameter mechanism design and sequential posted pricing. In *STOC*, pages 311–320. ACM, 2010.

[5] José R. Correa, Andrés Cristi, Boris Epstein, and José A. Soto. The two-sided game of googol and sample-based prophet inequalities. In *SODA*, pages 2066–2081. SIAM, 2020.

[6] José R. Correa, Andrés Cristi, Laurent Feuilloley, Tim Oosterwijk, and Alexandros Tsigonias-Dimitriadis. The secretary problem with independent sampling. In *SODA*, pages 2047–2058. SIAM, 2021.

[7] José R. Correa, Paul Dütting, Felix A. Fischer, and Kevin Schewior. Prophet inequalities for I.I.D. random variables from an unknown distribution. In *EC*, pages 3–17. ACM, 2019.

[8] José R. Correa, Paul Dütting, Felix A. Fischer, Kevin Schewior, and Bruno Ziliotto. Unknown I.I.D. prophets: Better bounds, streaming algorithms, and a new impossibility (extended abstract). In *ITCS*, volume 185 of *LIPIcs*, pages 86:1–86:1. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2021.

[9] José R. Correa, Patricio Foncea, Ruben Hoeksma, Tim Oosterwijk, and Tjark Vredeveld. Posted price mechanisms and optimal threshold strategies for random arrivals. *Math. Oper. Res.*, 46(4):1452–1478, 2021.

[10] José R. Correa, Patricio Foncea, Dana Pizarro, and Victor Verdugo. From pricing to prophets, and back! *Oper. Res. Lett.*, 47(1):25–29, 2019.

[11] Constantinos Daskalakis and Manolis Zampetakis. More revenue from two samples via factor revealing sdps. In *EC*, pages 257–272. ACM, 2020.

[12] Peerapong Dhangwatnotai, Tim Roughgarden, and Qiqi Yan. Revenue maximization with a single sample. *Games Econ. Behav.*, 91:318–333, 2015.

[13] Paul Dütting, Silvio Lattanzi, Renato Paes Leme, and Sergei Vassilvitskii. Secretaries with advice. In *EC*, pages 409–429. ACM, 2021.

[14] Tomer Ezra, Michal Feldman, Nick Gravin, and Zhihao Gavin Tang. Online stochastic max-weight matching: Prophet inequality for vertex and edge arrival models. In *EC*, pages 769–787. ACM, 2020.

[15] Michal Feldman, Nick Gravin, and Brendan Lucier. Combinatorial auctions via posted prices. In *SODA*, pages 123–135. SIAM, 2015.

[16] Moran Feldman, Ola Svensson, and Rico Zenklusen. Online contention resolution schemes. In *SODA*, pages 1014–1033. SIAM, 2016.

[17] Hu Fu, Nicole Immorlica, Brendan Lucier, and Philipp Strack. Randomization beats second price as a prior-independent auction. In *EC*, page 323. ACM, 2015.

[18] Nikolai Gravin and Hongao Wang. Prophet inequality for bipartite matching: Merits of being simple and non adaptive. In *EC*, pages 93–109. ACM, 2019.

[19] Chenghao Guo, Zhiyi Huang, Zhihao Gavin Tang, and Xinzhi Zhang. Generalizing complex hypotheses on product distributions: Auctions, prophet inequalities, and pandora's problem. In *COLT*, volume 134 of *Proceedings of Machine Learning Research*, pages 2248–2288. PMLR, 2021.

[20] Mohammad Taghi Hajiaghayi, Robert D. Kleinberg, and Tuomas Sandholm. Automated online mechanism design and prophet inequalities. In *AAAI*, pages 58–65. AAAI Press, 2007.

[21] Haim Kaplan, David Naori, and Danny Raz. Competitive analysis with a sample and the secretary problem. In *SODA*, pages 2082–2095. SIAM, 2020.

[22] Robert Kleinberg and S. Matthew Weinberg. Matroid prophet inequalities. In *STOC*, pages 123–136. ACM, 2012.

[23] Ulrich Krengel and Louis Sucheston. Semiamarts and finite values. *Bulletin of the American Mathematical Society*, 83(4):745–747, 1977.

[24] Ulrich Krengel and Louis Sucheston. On semiamarts, amarts, and processes with finite value. *Probability on Banach spaces*, 4:197–266, 1978.

[25] Aviad Rubinstein. Beyond matroids: secretary problem and prophet inequality with general constraints. In *STOC*, pages 324–332. ACM, 2016.

[26] Aviad Rubinstein, Jack Z. Wang, and S. Matthew Weinberg. Optimal single-choice prophet inequalities from samples. In *ITCS*, volume 151 of *LIPIcs*, pages 60:1–60:10. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2020.

[27] Ester Samuel-Cahn. Comparison of threshold stop rules and maximum for independent nonnegative random variables. *the Annals of Probability*, pages 1213–1216, 1984.